Energy harvesting and energy efficient control for variable impedance actuators

Wu, Fan

Awarding institution:
King's College London

Department of Engineering
King's College London
WC2R 2LS London
United Kingdom

Energy Harvesting and Energy Efficient Control for Variable Impedance Actuators

---

Fan Wu

Primary Supervisor:
Dr Matthew J. Howard

Secondary Supervisor:
Dr Daniele Magazzani

A thesis submitted as part of the requirements for the degree of Doctor of Philosophy in Robotics.

Abstract

The thesis aims at improving energy efficiency of compliant actuators from a new perspective, by incorporating energy harvesting technology. Energy efficiency is a crucial issue towards long-term deployment of compliant robots in the real world. In the context of variable impedance actuators (VIAs), one of the main focuses has been on improving energy efficiency through reduction of energy consumption. However, the harvesting of dissipated energy in such systems remains underexplored.

The first part of the thesis proposes a novel variable damping module design enabling energy regeneration in VIAs by exploiting the regenerative braking effect of DC motors. The proposed damping module uses four switches to combine regenerative and dynamic braking, in a hybrid approach that enables energy regeneration without reduction in the range of damping achievable. Numerical simulations and a physical experiment were conducted in which the proposed module shows an optimal trade-off between task-performance and energy efficiency.

In addition, to investigate the role of variable regenerative damping in terms of energy efficiency of long-term operation, experiments are reported in which the VIA, equipped with the proposed damping module, performs sequential reaching to a series of stochastic targets. The results indicate that the combination of variable stiffness and variable regenerative damping results in a 25% performance improvement on metrics incorporating reaching accuracy, settling time, energy consumption and regeneration over comparable schemes where either stiffness or damping are fixed.

Next, to consider non-periodic but sequential tasks, instead of optimizing the whole problem, we propose to treat the movement as a sequential combination of sub-movements and establish a bi-level framework. Low-level optimal control of

subtasks is encapsulated in an outer loop of policy improvement to enable high-level optimization w.r.t. realistic energy cost estimators. The proposed method is able to (i) exploit variable physical impedance, (ii) tune efficiency balance via control cost weighting, (iii) optimize movement timing. The effectiveness is verified by two consecutive reaching tasks, and significant energy saving is demonstrated by simulation and measurement on hardware.

Acknowledgements

First of all, I would like to thank my primary supervisor Dr Matthew Howard for the support he has given, without whom I would not have had the chance to delve into the field of Robotics Research. My research ability and academic writing has been dramatically improved with his invaluable advice and guidance. I would also like to thank the lab members for their support throughout my research. Dr Brendan Michael and Dr Jeevan Manavalan helped start the work at King's. I would like to thank Dr Yu-Chen Zhao for his inspiring discussions with me.

I would like to express my gratitude to James Trotter for his technical support, and for his patience with my naive questions regarding mechatronics development. I would also like to thank Jon West for his help with electronics improvement.

I am really grateful to the donator of Professor Sir Richard Trainor PhD Scholarship for the financial support for this study. Without it my PhD would not be able to start.

Also, I would like to acknowledge the people in Centre for Robotics Research, Department of Engineering and Department of Informatics, Dr Teng Sun, Jian Hu, and Dr Long Chen, for their discussions, encouragements, and friendship. I would like to thank Dr Xin-Sheng Zhang for that he shared a lot of knowledge, thoughts and opinions (and also lots of interesting stories) that helped me kick-started my life as a researcher.

I would like to thank my parents for their incredible support during my PhD, and for allowing me to do whatever I wish. During the time when I was composing this thesis in my hometown with them, we encountered the Coronavirus outbreak and have spent a special long holiday together under city lockdown.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(Fan Wu)

Contents

List of Figures

List of Tables

Abbreviations

BBO: Black-Box Optimization

CMA: Covariance Matrix Adaptation

DC: Direct Current

DDP: Differential Dynamic Programming

DMP: Dynamic Movement Primitives

DOF: Degree of Freedom

EF: Efficient Frontier

EIDC: Extended Inverse Dynamics Controller

EMF: Electromotive Force

EMG: Electromyography

EP: Equilibrium Position

ER: Electro-rheological

ES: Evolution Strategy

HJB: Hamilton-Jacobi-Bellman

ILQR: Iterative Linear Quadratic Regulator

IOC: Inverse Optimal Control

IRL: Inverse Reinforcement Learning

IR: Industrial Robot

KERS: Kinetic Energy Recovery System

MACCEPA: Mechanically Adjustable Compliant and Controllable Equilibrium Position Actuator

MR: Magneto-rheological

OC: Optimal Control

OCP: Optimal Control Problem

ODE: Ordinary Differential Equation

PEA: Parallel Elastic Actuator
PMP: Pontryagin's Maximum Principle
PWM: Pulse Width Modulation
RK4: Runge-Kutta Fourth
RL: Reinforcement learning
ROS: Robot Operating System
SEA: Series Elastic Actuator
SSM: State-Space Model
TIDC: Tracking with Inverse Dynamics Controller
VD: Variable Damping
VIA: Variable Impedance Actuator
VSA: Variable Stiffness Actuator
VSFD: Variable Stiffness and Fixed Damping
VSVD: Variable Stiffness and Variable Damping

Chapter 1

Introduction

This is an introductory chapter of the thesis to introduce the background, research aims, and outline the contents and contributions.

1.1 Background

Variable impedance actuators (VIAs) are believed to be the key for incorporating morphological intelligence, which is found in biological systems, into the next generation of robots (Vanderborght et al. (2012)). These actuators have the ability to vary the mechanical impedance (stiffness and damping), whose subgroup variable stiffness actuators (VSAs) do not have a damping modulator. Traditional stiff robotics are well-developed for accurate tracking of a desired trajectory with high gain and high bandwidth control. The growing demands in service robots, assistive devices, robotic surgery, *etc.*, cannot be met by the traditional robots driven by stiff actuators which are unable to interact safely and robustly in human environment.

To circumvent the limitation of stiff actuation, compliant actuators[1] have been in development for decades to implement direct mechanical impedance modulation (Hogan and Buerger (2004)). In the meanwhile, the physical compliance incorporated in VSAs (*e.g.,* using elastic components such as springs) enables

[1]Compliant robots can have rigid links or soft links. The compliant actuators discussed in this thesis are assumed to be attached to rigid-body robots. The soft link or variable stiffness link (VSL) robots are beyond the scope of this thesis.

energy storage, which can be used to (i) absorb external energy introduced into the system (*e.g.,* from collisions) to enhance safety, and (ii) amplify output power by releasing stored energy as and when required by the task (Grebenstein et al. (2011)). This property makes compliant actuation a promising approach to not only achieve inherent safety and human-level functional performance for dynamic interaction, but also improve the energy efficiency of robots towards the level of biological systems.

Different from human muscles, of which the mechanical stiffness and damping are coupled, mechatronic systems can decouple the two via design and modulate them independently. This fosters a direction in which robots driven by VIAs can exploit the independent damping modulation rather than simply imitating human movements as shown by Zhao et al. (2018), and possibly outperform humans. Much research effort has recently gone into the design of *variable physical damping* actuation, based on different principles of damping force generation (see Tagliamonte et al. (2012); Vanderborght et al. (2013) for a review). Variable physical damping has proven to be necessary to achieve better task performance, for example, in eliminating undesired oscillations caused by the elastic elements of VSAs (Laffranchi et al. (2012a,b)). It has also been demonstrated that variable physical damping plays an important role in terms of energy efficiency for actuators that are required to operate at different frequencies, to optimally exploit the natural dynamics (Laffranchi et al. (2012a)).

However, the importance of variable physical damping may be underestimated, because the potential to *harvest energy dissipated by damping* has so far received little attention. While recent studies of compliant actuation design and control represent important advances in terms of improving the efficiency of *energy consumption* in VIAs, the field has not seen research efforts exploring the approach of *energy harvesting*.

To address this, the thesis proposes to utilize the energy regeneration capabilities of electromechanical systems to develop a damping module that can exert a controllable braking effect and recover kinetic energy into electricity. To the author's knowledge, this is the first attempt to show the potential of higher energy efficiency of VIAs by incorporating energy recovery into the damping module.

1.2 Motivation

1.2.1 Energy saving

As introduced above, the first main motivation is to fill the energy saving gap with current research addressing energy efficiency of VIAs. In industrial robotics and process automation, energy efficiency has been a crucial issue to make manufacturing sustainable and competitive. The mass adoption of energy intensive industrial robots (IR) compromises the sustainability of manufacturing (Pellicciari et al. (2015)). For example, 8% of the total electrical energy usage in production processes of automotive industries is consumed by industrial robots (Paryanto et al. (2015)). It has become increasingly crucial for robotized factories to optimize operating costs and reduce ecological footprint such as CO_2 emission. The European Commission has set up the AREUS project aiming at making factories *smart and green*. For ecological and economic reasons, this motivates research on reducing the energy cost of IRs. For more details of the attempts to reduce energy consumption of IRs, we refer the reader to Carabin et al. (2017). Furthermore, with the extensive deployment of compliant robots expected in the near-future for human-robot collaboration, medical and civil services, this imperative to save energy is likely to become even more critical.

An important factor that hinders the long-term deployment of current mobile and humanoid robots is their low energy efficiency. *The most efficient humanoid's locomotive energy cost still far exceeds human and animals (Kashiri et al. (2018)).* Compliant robotics is a promising solution towards energy efficient human-like movement. The recently developed bipedal humanoid robot DURUS (Reher et al. (2016)) incorporating compliant ankles provably shows 70% less energy expenses compared to its counterpart ATLAS developed by Boston Dynamics. A biarticular compliance mechanism mimicking human limbs has demonstrated a human-level efficient squatting motion on a prototype developed by Roozing et al. (2019). Yet energy efficiency of general purpose actuators with variable compliance still has much potential for improvement to make significant impact on real world applications.

1.2.2 Energy efficient control of complex movement

Understanding how to improve the energy efficiency of compliant robotic systems is a scientific question related to energy economy of biomechanical systems. Biological springs, like tendons and various elastic elements in muscles embedded in human and animals, make them highly efficient runners and jumpers (Roberts (2016)). The field of compliant robotics has seen a series of successful developments of energy efficient robots with elastic joints or springy legs for (i) robotic locomotion, (ii) explosive movements (Wolf and Hirzinger (2008); Braun et al. (2013)), and (iii) cyclic manipulation tasks (Matsusaka et al. (2016); Haddadin et al. (2018)). In these examples, the energy storage of physical compliance is exploited to generate natural and efficient motion.

However, many tasks in unstructured environments are not periodic. For instance, the objects to be picked and placed may be located at random positions. The task given to a robot may consist of a sequence of different types of actions, such like "reach a cup, grasp it, and pour the water". These non-periodic but *sequential* tasks more commonly involve upper limbs and are complicated by the greater diversity of the tasks. It is of great interest to apply the same principle to robotic manipulators such that compliant robots can behave in a human-like energy efficient way for a wide variety of tasks. How to design a generalizable method for energy-optimal trajectory generation of robots driven by compliant actuators is still an open question.

To answer this question, the thesis will discuss how to realize the skill of compliance (Okada et al. (2002))—the ability to exploit elastic energy to improve energy efficiency—for sequential movements. We believe this ability is key for compliant robots to display skilled and efficient human-like movements. The approach proposed will employ a hierarchical structure that treats the high-level optimization of 'compliance skill' as a model-free reinforcement learning problem and generate the low-level controllers using an optimal control method. To the author's knowledge, no existing approach has been demonstrated to address (i) exploiting variable physical impedance, (ii) optimally tune control effort weighting, and (iii) optimizing temporal features, for VIA actuated sequential movements.

1.3 Aims and Objectives

In summary, to address the issues and answer the questions raised above, the thesis has the following aims and objectives:

1. Design a regenerative damping module that can provide a controllable damping effect and harvest energy from bi-directional joint movements;
2. Implement and evaluate the regenerative damping design on a physical robot to show significance of energy efficiency improvement;
3. Extend the optimal control method to address sequential movements, with the help of model-free reinforcement learning and an efficient evolutionary optimization method.

1.4 Publications

Below is the list of publications or submissions by the author when pursuing his PhD. Publications I-III directly contribute the main content of this thesis. Items IV, V are relevant in our discussions in Chapter 2, 3.

I **Fan Wu** and Matthew Howard, Exploiting Variable Impedance for Energy Efficient Sequential Movements, 2020, preprint, arXiv:2002.12075.

II **Fan Wu** and Matthew Howard, Energy regenerative damping in variable impedance actuators for long-term robotic deployment, IEEE Transactions on Robotics (T-RO), 2020, doi: 10.1109/TRO.2020.2998641.

III **Fan Wu** and Matthew Howard, A Hybrid Dynamic-regenerative Damping Scheme for Energy Regeneration in Variable Impedance Actuators. IEEE International Conference on Robotics and Automation (ICRA), 2018.

IV YuChen Zhao, Aran Sena, **Fan Wu**, Matthew Howard. A Framework for Teaching Impedance Behaviours by Combining Human and Robot 'Best Practice', IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2018.

V Samuel Pitou, **Fan Wu**, Ali Shafti, Brendan Michael, Riaan Stopforth, Matthew Howard, Embroidered Electrodes for Control of Affordable Myographic Prostheses. IEEE International Conference on Robotics and Automation (ICRA), 2018.

1.5 Thesis outline

The thesis is structured as follows.

In **Chapter 2**, we review the research efforts in the literature so far addressing the energy efficiency of VIAs.

Contributions:

- An analysis based on the power flow of VIAs is accounted, which summarizes the previous works and points out an alternative direction of improving energy efficiency by harvesting the energy dissipated by the damper.

Relevant publications: II

In **Chapter 3**, we present the modelling and optimal control of VIAs. Then we discuss the previous work and an alternative approach to address the energy efficiency which is never explored before. Furthermore, we discuss the limitations of optimal control especially when it comes to complex non-periodic movements.

Contributions:

- A novel tracking controller that naturally resolves actuation redundancy for VIAs is provided for the first time.

Relevant publications: Publication I

In **Chapter 4**, we propose a novel variable regenerative damping module design enabling energy regeneration in VIAs by exploiting the regenerative braking effect of DC motors. The proposed damping module uses four switches to combine regenerative and dynamic braking, in a hybrid approach that enables energy regeneration without reduction in the range of damping achievable. Numerical simulations and a physical experiment are presented in which the proposed module shows an optimal trade-off between task-performance and energy efficiency.

Contributions:

- A novel regenerative damping module design capable of harvesting energy from bidirectional movements, without loss of maximum damping effect available.
- Experiment demonstrates a linear relationship between damping effect and

control which is in good agreement with the model prediction. This makes it easy to be integrated into model-based on optimal control.
- Evaluation of optimal control as a means for dealing with the energy/task performance trade-off. Simulations show the potential of energy efficiency improvement.

Relevant publications: Publication III

In **Chapter 5**, we present experimental work conducted to evaluate the proposed regenerative damping in the context of long-term operation. To investigate the role of variable regenerative damping in terms of energy efficiency in the context of long-term operation, experiments are reported in which the VIA, equipped with the proposed damping module, performs sequential reaching to a series of stochastic targets.

Contributions:
- A long-term operation setting with stochastic reaching target and evaluation metric including accuracy, settling time, energy consumption and regeneration. By which the energy efficiency of a general purpose VIA can be evaluated and analysed.
- Experimental data verified that exploiting variable stiffness and variable damping simultaneously is preferred.
- New view of the role of damping as a kinetic energy recovery system to complement the elastic energy storage mechanism. When natural dynamics can not be exploited as in pure rhythmic movement, energy regeneration comes to recover the energy for discrete movement. With a variable stiffness mechanism, over-exploitation of damping, which results in dynamic braking, can be avoided to make sure the regeneration works as desired.

Relevant publications: Publication II

In **Chapter 6**, we consider sequential movements in the context of a robot driven by compliant actuators. Sequential movements can be viewed as an abstraction of a broad class of complex movements including the consecutive point-to-point reaching used in Chapter 5. We integrate optimal control and reinforcement learning to exploit variable impedance in the sequential context. Meanwhile,

control cost weighting and temporal features can be optimized.

Contributions:

- Establish a hierarchical framework that optimizes high-level motion plan in an iterative policy learning fashion, taking account of variable (physical) impedance exploitation, and encapsulates optimal control and or tracking control at the low-level.
- Experiments evaluated to demonstrate the effectiveness of proposed framework to optimize performance-cost trade-off, with significant energy saving of sequential movements reported.
- The approach takes general formalization and has potential to be extended to other compliant or soft robots and consider more complex sequential movements in the real world.

Relevant publications: Publication I

Finally, **Chapter 7** summarizes conclusions and suggests directions for future work.

Chapter 2

Variable Impedance Actuators

This chapter provides a literature review of VIAs from the perspective of energy efficiency. We will briefly introduce the concepts of VIAs and present a basic model. Based on the model and analysis of its power flow, we categorize previous works and point out an alternative way to improve energy efficiency by energy regeneration.

2.1 Introduction

Industrial robots are typically designed by the traditional premise — *"the stiffer the better"* (Salisbury et al. (1991)). They are driven by stiff actuators equipped with high ratio gearboxes and controlled with high gain to ensure positioning accuracy and stability. In contrast, human motor systems made by muscles, tendons and ligaments perform voluntary or skilful movements with significant ease and efficiency, in spite of slow[1] and noisy neural signals (Faisal et al. (2008)). Regarding the energy efficiency of locomotion measured by the Cost of Transport[2] (CoT), an adult human with average weight has a CoT of 0.2 when walking

[1]The nerve conduction velocity is up to 120m/s (Siegel and Sapru (2005)), which is much slower than the speed of electricity.

[2]Cost of transport is a dimensionless quantity that allows to compare the energy efficiency of locomotion among dissimilar animals and modes of transportation. It is defined as $\mathrm{CoT} := \frac{E}{mgd}$, where E is the total energy consumption for travelling a distance d, m is the mass and g represents the standard gravity.

(Tucker (1975)), while the CoT of the famous humanoid ASIMO[3] is 2.3, which is more than ten times greater than that of a human. Bauml et al. (2010) presented an impressive ball catching experiment on the DLR-LWR-III robot, there the robot caught an $80\,\mathrm{g}$ flying ball at a speed of $25\,\mathrm{km/h}$. In comparison, a human handball goalkeeper can withstand the impact of a $480\,\mathrm{g}$ ball at $120\,\mathrm{km/h}$, which means 100 times more energy.

Why does the performance of biological motor systems, in terms of energy efficiency and impact tolerance according to the above two cases, far exceed their robotic counterparts? It is in part due to their inherent elasticity and energy-storage capability. Exploitation of this natural property results in efficient and robust behaviours in an unknown and dynamic environment without the need of huge computation. This has been viewed as a kind of embodied intelligence by an increasing number of researchers (Pfeifer and Bongard (2007) & Vanderborght et al. (2012)).

Elasticity is common and of great importance in various aspects of animal movements (Alexander (1988)). Biological springs — tendons, ligaments, and various elastic elements in muscles — embedded in human and animals are what make them highly efficient runners and jumpers (Roberts (2016)). Utilizing elastic energy storage and recoil, which is associated with optimizing muscular stiffness and transition timing, is a crucial skill that can be practised and improved for many other athletic activities, not limited to locomotion (Wilson and Flanagan (2008)). Burdet et al. (2001) and Franklin et al. (2008) confirmed by experimental studies that humans learn to stabilize unstable dynamics by modulating the joint impedance, which indicates that humans use impedance control (Hogan (1985)) and exploit the natural dynamics rather than overcome it (Radulescu (2016)).

The idea of stably controlling interaction behaviours by regulating impedance[4] starts with seminal work by Hogan (1985). Following this idea, active compliance control is used on stiff robots, *e.g.,* the KUKA/DLR lightweight arm (Albu-Schäffer et al. (2011)) to mimic the spring-like behaviour for safe interaction and collaboration (Ikeura and Inooka (1995)). However, when unexpected collision

[3] According to Sakagami et al. (2002), ASIMO has a mass of $54\,\mathrm{kg}$, costs $1.8\,\mathrm{kW}$ for a walking speed of $1.5\,\mathrm{m/s}$

[4] Impedance includes stiffness and damping.

occurs, due to the limited control bandwidth and stiff actuation, the shock energy has no way to be absorbed to protect both human and robot (Haddadin et al. (2009)).

The most well-known general purpose compliant actuator is the Series Elastic Actuator (SEA) proposed by Pratt and Williamson (1995). Prior to that, the earliest works explored incorporating physical compliance emerged by the 1980s. For instance, a common issue in the early days of IRs was the difficulty with accurate insertion during assembly in presence of misalignment of components. In contrast to relying on precise force sensing, McCallion et al. (1979) proposed a simple passive device with physical compliance, *i.e.,* springs, to complete closely-fitting insertion of peg-in-hole by adapting to the interaction rather than overcoming it. Springy leg was adopted into the design of legged robots by Raibert (1986) and enhanced dynamic balance was achieved.

The need for safety and energy efficiency has motivated lots of research efforts devoted to compliant actuation[5] over two decades, with increasing emphases on realizing variable and controllable stiffness and damping. A variety of designs have been proposed based on manifold cognate principles and found applications in (i) robotic locomotion (Yamaguchi et al. (1998); Robinson et al. (1999); Pratt (2000); Collins and Ruina (2005); Van Ham et al. (2007); Vanderborght et al. (2011); Hutter et al. (2013); Reher et al. (2016); Spröwitz et al. (2018)), (ii) protheses and exoskeleton (Paluska and Herr (2006); Au et al. (2007); Cherelle et al. (2012); Yu et al. (2013); Grimmer et al. (2014)), and (iii) physical human-robot interaction (Kong et al. (2009); Albu-Schäffer et al. (2011); Grebenstein et al. (2011); Goris et al. (2011); Laffranchi et al. (2012c); Wolf et al. (2015))

2.2 Variable Impedance Actuators

The first general purpose series elastic actuator was proposed by Pratt and Williamson (1995) to reduce the reflected inertia and enhance shock tolerance of the actuation system. As shown in Figure 2.1(a), the spring was serially placed between the load and the motor (with a gearbox). By absorbing the energy introduced by

[5]Compliance is the opposite of stiffness, so the variable stiffness and adjustable compliance are often refer to the same concept.

shock, the spring protects the gearbox from breaking its mechanical limits and the impact tolerance of the whole actuator is increased. Another benefit is that the force control problem is converted into a position control problem, which alleviates the need for a precise force sensor. Today SEA is still widely adopted in recent developments of legged robots (Hutter et al. (2013); Roozing et al. (2019)).

The growing field of compliant robotics has seen a variety of innovative designs for adjustable stiffness based on the principles of (i) antagonistic springs, (ii) variable transmission ratio, and (iii) pretension adjustment of single spring.

Antagonistic springs

One antagonistic spring structure is the traditional but popular arrangement shown below (Figure 2.1(b)). It is a biologically-inspired mechanism first implemented in the McKibben Artificial Muscle via pneumatic actuation. To achieve variable stiffness the springs have to be non-linear. Apart from using a real non-linear spring, this can be realized by geometric non-linearity. Examples are, to name a few, the Quasi-Antagonistic Joint (QA-Joint) (Eiberger et al. (2010)), cross-coupled with a third spring VSA-II (Schiavi et al. (2008)), bidirectional coupled springs VSA-cube (Catalano et al. (2011)). Antagonistic springs are being used to build musculoskeletal robotic systems such as Myorobotics (Maufroy et al. (2014)), Anthrob (Jäntsch et al. (2013, 2015)) and the humanoid Roboy (Richter et al. (2016)).

Variable transmission ratio

Most variable transmission ratio mechanisms are implemented by varying lever length between the load and spring. As shown in Figure 2.1(c), by changing the pivot point, the ratio between the deflection and spring force varies. A series of works have been carried out by Jafari et al. (2013, 2014, 2015), which led to energy efficient actuators AwAS and AwAS-II. A similar principle was used by Groothuis et al. (2014) for developing vsaUT-II. One advantage is that a large stiffness range can be realized via this principle which make it desirable when a high stiffness is needed for precise positioning. Secondly, the stiffness modulation is energetically efficient since the movement direction of the pivot point is perpen-

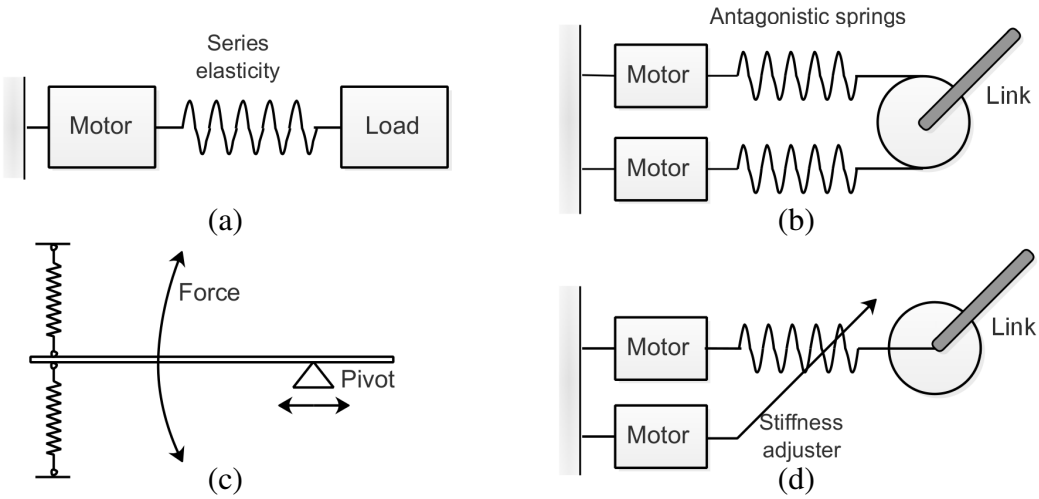


Figure 2.1: Block diagrams of different ways to incorporate elasticity: (a) series elastic actuators, (b) antagonistic springs, (c) variable stiffness adjusted by controlling transmission ratio, and (d) single spring with adjustable preload.

dicular to the force generated by the springs (more rigorously speaking, when the arm is at equilibrium position).

Adjustment of spring pretension

Figure 2.1(d) illustrates a single spring connected to the joint in series, usually through a non-linear geometry, and another motor is used to adjust the stiffness. A lever arm is used in MACCEPA (Van Ham et al. (2007)) to generate non-linearity and the pretension of a linear spring is varied to adjust the output stiffness. In MACCEPA 2.0 the lever arm is replaced with a cam disk (Van Ham et al. (2009)). To improve the compactness, Furnemont et al. (2015) redesigned it with the use of a torsional spiral spring, which was easy to be mounted around the drive axis. The VS-joint (Wolf and Hirzinger (2008)) and FS-joint (Wolf et al. (2011)) are also categorized into this type. There, instead of modulating the spring length directly, the stiffness adjuster controls a preload to push a cam disk and the spring is compressed.

These are common ways to incorporate variable stiffness for actuating a single joint. Note that, in contrast to the SEA, parallel elastic actuator (PEA) places the spring between the load and base in the opposite of the motor. It can be seen as an asymmetric antagonistic design using a stiff motor on one side and a passive

spring on the other side. An analysis and comparison of the energy consumption between SEA and PEA was provided by Verstraten et al. (2016). The PEA has been involved in more complex mechanisms for efficient locomotion recently. For example, a *(SEA + PEA)* method was employed by Grimmer et al. (2012) and showed that peak power requirement was further reduced. Inspired by the articulated muscle arrangement in human limbs, a series-parallel mechanism is being developed aiming at human-like efficient locomotion (Roozing et al. (2019)).

2.2.1 Variable physical damping

VIAs are those VSAs with a dedicated damping module. The damper can be arranged in the mechanism in different ways as illustrated in Figure 2.2. Variable physical damping has proven to be necessary to achieve better task performance, for example, in eliminating undesired oscillations caused by the elastic elements of VSAs (Laffranchi et al. (2012d, 2013, 2012b)). It has also been demonstrated that variable physical damping plays an important role in terms of energy efficiency for actuators that are required to operate at different frequencies, to optimally exploit the natural dynamics (Laffranchi et al. (2012a)). Independent control of stiffness and damping is where robots differ from humans. Human muscles' co-contraction increase stiffness and damping in a way that the damping ratio nearly unchanged (Milner and Cloutier (1998)). By contrast, for VIAs damping is decoupled from stiffness and thus can be exploited. The idea of exploiting damping control to outperform human has been discussed in Zhao et al. (2018), where an active impedance controlled robot was used. Equipped with variable physical damping, compliant robots have the potential to learn kinematic movement and stiffness profile (Howard et al. (2013)) from human demonstration — by electromyography (EMG) sensing (via *e.g.,* Pitou et al. (2018)) — and optimize the damping profile for maximum performance and or minimizing energy (Zhao et al. (2018)).

Principles of variable physical damping

Variable (controllable) damping has been designed based on different principles. As reviewed by Vanderborght et al. (2013); Tagliamonte et al. (2012) most vari-

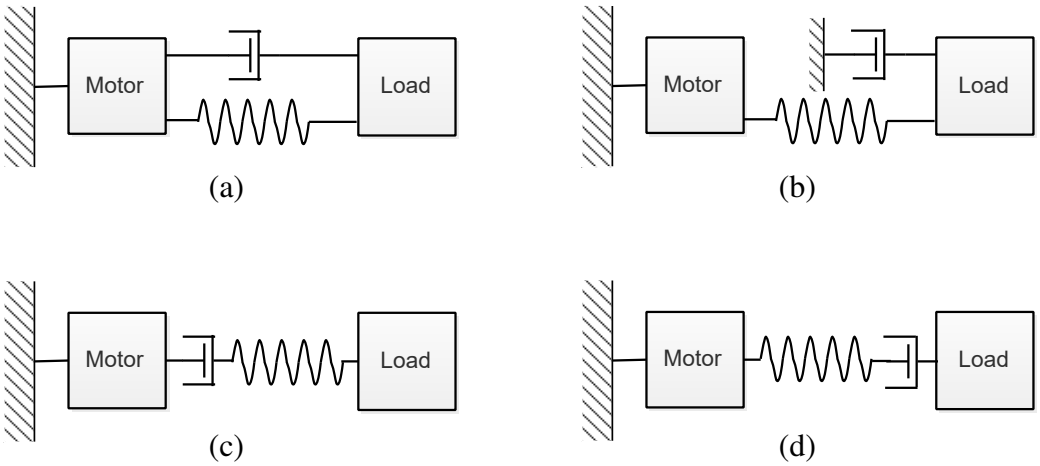


Figure 2.2: Block diagrams of different ways to incorporate damping: (a) pure parallel, (b) external parallel, (c) serial damping first, and (d) serial spring first. The damping element can be introduced through a pure damper or with visco-elastic elements (Vanderborght et al. (2013)).

able physical damping design in robotics are based on: (i) friction, (ii) Electrorheological (ER) and magnetorheological (MR), (iii) fluid dynamics, and (iv) eddy current. A frictional damper applies a normal force to the contact surface as implemented in Morita and Sugano (1995). A piezoelectric-based frictional damper was designed by Laffranchi et al. (2010) for a SEA. Electro-rheological (ER) and magneto-rheological (MR) dampers use electrical or magnetic fields to change the viscosity of fluid. Both frictional and ER/MR dampers are commonly used in vehicles, However, they both present high hysteresis as pointed by Guglielmino et al. (2005). Fluid dynamics damper uses a valve to adjust the fluid viscosity *e.g.,* in Catalano et al. (2012). Eddy current is generated when conductive material is moving in a magnetic field and an electromotive force (EMF) is created to counter the movement. This is a common effect in electric motors and thus a DC motor can be easily implemented as a damper as demonstrated in Radulescu et al. (2012).

Combinations of elastic and damping elements

There are various possible system topologies resulting from different combinations of elastic and damping elements. Figure 2.2 illustrates four basic cases by

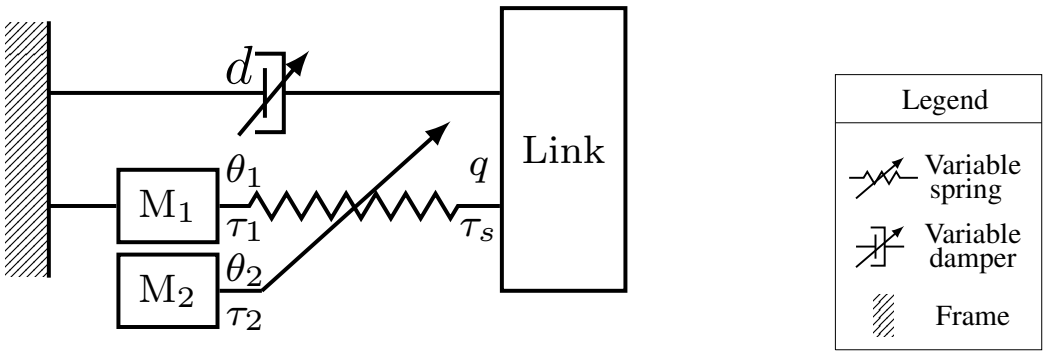


Figure 2.3: Schematic diagrams of a VIA represented by a mass-spring-damper model. The VIA has compliance incorporated in series between the motor and link

assuming that: (1) a load (or a link) is actuated by one motor which is connected to a reference frame, (2) an elastic element is connected between the motor and the load, and (3) the damping is incorporated either via a pure damper or viscoelasticity.

The first case in Figure 2.2(a) represents an elastic element and a damper arranged in parallel between the motor and the load. The link side dynamics of this system is described by

$$m\ddot{q} + d(\dot{q} - \dot{\theta}) + k(q - \theta) = 0, \tag{2.1}$$

where m denotes the link inertia, q the link position, θ the motor position, d, k are the damping and stiffness.

The second possibility shown by Figure 2.2(b) shows a damper connected to the reference frame, which results in a system dynamics written as

$$m\ddot{q} + d\dot{q} + k(q - \theta) = 0. \tag{2.2}$$

It can been seen that in dynamics equation (2.1) the damping force is proportional to the relative velocity $(\dot{q} - \dot{\theta})$, whereas in (2.2) the damping force $d\dot{q}$ is unrelated to the motor velocity.

The serial combinations shown in Figure 2.2(c) and (d) result in two similar but different system dynamics. When a damper is connected between the motor

and the spring (Figure 2.2(c)), the system dynamics is written as

$$\begin{aligned} m\ddot{q} + k(q - q_c) &= 0 \\ d(\dot{q}_c - \dot{\theta}) + k(q_c - q) &= 0, \end{aligned} \tag{2.3}$$

where q_c is the position at the point connecting the damper and spring. When the spring is connected first to the motor (Figure 2.2(d)), the system dynamics becomes

$$\begin{aligned} m\ddot{q} + d(\dot{q} - \dot{q}_c) &= 0 \\ d(\dot{q}_c - \dot{q}) + k(q_c - \theta) &= 0. \end{aligned} \tag{2.4}$$

The main advantage that scheme (b) has over the other three is that it allows the damping element to be designed as a modular unit and attached externally to the variable stiffness mechanism. A detailed discussion about the advantages and disadvantages of each scheme was given by Vanderborght et al. (2013) .

The thesis focuses on VIA mechanisms that can be represented by the mass-spring-damper model depicted in Figure 2.3, where the variable stiffness mechanism is based on Figure 2.1(d). This model covers a lot of actuator designs whose variable stiffness mechanism is arranged in a series configuration. The link, whose position is denoted by q, is connected in series to a motor and a variable spring. The equilibrium position is controlled by the motor M_1 and the stiffness is modulated by another motor M_2, whose positions are denoted by θ_1 and θ_2, respectively. The torques exerted by the spring on the link, M_1, and M_2 are τ_s, τ_1, τ_2. The damper in this system is arranged between the link and the base as shown in Figure 2.2(b), and the damping d is independently controllable.

2.3 Energy Efficient Compliant Actuation

The problem of energy efficiency in compliant robotic systems has been addressed via different approaches from the perspectives of control or design, which can be mainly categorized into studies that (i) look at *exploiting energy storage* in periodic or discrete movements, or (ii) focus on *reducing energy consumption* via the mechanical design. In the following subsections, the basic concept of power flow of VIAs is introduced as a means to understand these approaches, and point

out a third possible way to improve the energy efficiency by regenerating energy dissipated by damping (Wu and Howard (2020a)). This is followed by an account of the theory about regenerative braking of electric motors.

We present here a basic power flow model to support the review of literature from an energy perspective. The corresponding power flow of the robotic system driven by a VIA defined above (see Figure 2.3) is shown in Figure 2.5. A power source is assumed to supply the motors $\mathrm{M}_1, \mathrm{M}_2$. The elastic element, (*i.e.,* spring) can be viewed as an energy tank in the actuator that stores potential energy E_s. In

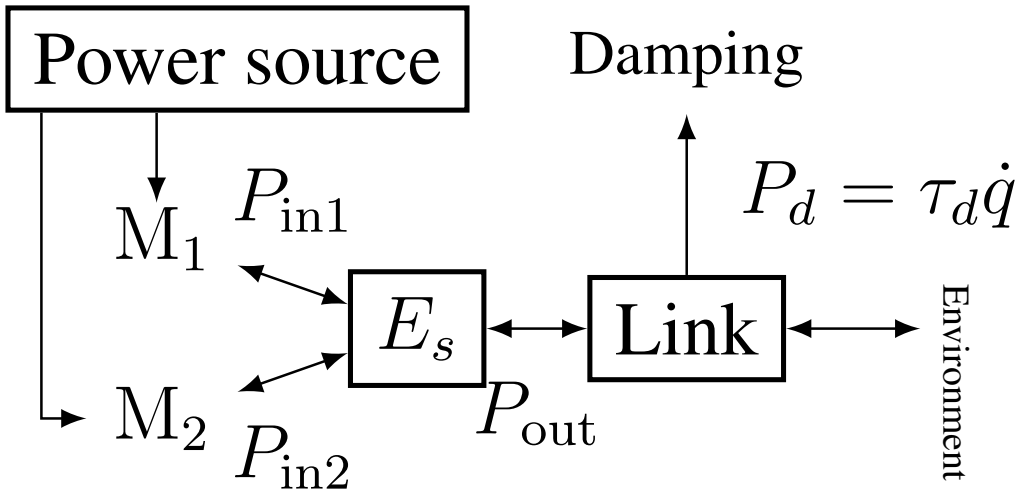


Figure 2.4: Schematic diagrams of the corresponding power flow of a VIA. Much energy is dissipated via damping when elastic energy storage is not possible to be exploited.

general, the power flow of this element can be represented in the form of power conversion $P_{\mathrm{in}} = P_{\mathrm{out}} + \dot{E}_s$, where $P_{\mathrm{in}}, P_{\mathrm{out}}$ is the power drained and delivered by the compliant actuation module, respectively, and $\dot{E}_s$ is the rate of change of energy stored. As shown in the diagram, in the types of VIAs considered in this

paper, P_{in} consists of power input from the two motors (P_{in1} and P_{in2}), hence

$$\begin{aligned}\dot{E}_s &= \frac{\partial E_s}{\partial q}\dot{q} + \frac{\partial E_s}{\partial \theta_1}\dot{\theta}_1 + \frac{\partial E_s}{\partial \theta_2}\dot{\theta}_2 \\ &= -\tau_s\dot{q} - \tau_1\dot{\theta}_1 - \tau_2\dot{\theta}_2 \\ &= -P_{\text{out}} + \underbrace{P_{\text{in1}} + P_{\text{in2}}}_{P_{\text{in}}}. \end{aligned} \tag{2.5}$$

P_{out} is bi-directional which means that the elastic element can deliver energy to, or receive energy from, the link that, in turn, exchanges energy with the environment via interaction.

Regulating the energy flow around the elastic element, as governed by (2.5), is one of the keys to improving the energy efficiency of compliant actuators. The majority of prior work in this area has focused on this issue, essentially prioritizing the problem of *energy consumption*. In this, two broad categories of approach can be identified: (i) exploiting energy storage E_s, and (ii) reducing energy cost of stiffness modulation P_{in2}.

2.3.1 Exploiting energy storage

The energy storage capability introduced by the elastic elements of VIAs brings not only the advantage of impact tolerance, but also the possibility of improving energy efficiency by reusing the stored energy. The power flow between the elastic element and output link is bidirectional, which means that energy can be stored in the elastic element, *e.g.,* by absorbing impact energy during interaction with the environment. It can also build up the potential energy by receiving the power from motors and then releasing at the best timing to amplify the output power. Energy storage occurs when $P_{\text{in}} > P_{\text{out}}$, and release occurs when $\dot{E}_s < 0$ contributing positive output power P_{out}. One of the appealing features of VIAs is that they can build up a reserve of energy in the elastic element by receiving power from motors, or through interactions with the environment, and time its release according to task demands.

Periodic movement

Following this idea, numerous mechanisms and control schemes for physically compliant actuation have been developed considering problems of periodic movements (*e.g.,* walking Collins and Ruina (2005); Vanderborght et al. (2006); VanderBorght et al. (2008); Hutter et al. (2013), lower-limb prosthetics Au et al. (2007) and cyclic manipulations Haddadin et al. (2011a); Lakatos et al. (2014); Velasco et al. (2015); Matsusaka et al. (2016); Haddadin et al. (2018)) and discrete movements (*e.g.,* throwing Braun et al. (2012); Ozparpucu and Haddadin (2013)).

While the use of springy legs has a long history and seminal work can be dated back to the 1980s (Raibert (1986)), Collins (Collins et al. (2005)) was the first to successfully demonstrate human-level efficiency on a passive-dynamic walker. One of the keys to its success was the implementation of compliant ankles based on findings from biomechanics about human walking. Vanderborght (Vanderborght et al. (2006); VanderBorght et al. (2008)) proposed to exploit the natural dynamics by fitting the compliance of the actuator to the natural compliance of the desired trajectory. The principle is followed by recent development of robotic locomotion driven by compliant actuators (Hutter et al. (2013)). However, new challenges, such as dynamic environments, uncertainties and the complexity of tasks facing robots running in the real world require more advanced tools like optimization-based numerical solutions to seek energy efficient control. In a recent experimental study, Baček et al. (2020) reported energy savings up to 50% via varying mechanical stiffness on a knee-joint actuator. Various studies comparing series elasticity with parallel elasticity (Grimmer et al. (2012); Verstraten et al. (2016); Kashiri et al. (2018)) suggest that further improvement of energy efficiency in locomotion can be achieved by integrating parallel elasticity for potential energy balancing (Roozing et al. (2016)). In this line of research, a bi-articulated spring arrangement has been utilized by biological inspiration and significantly improved the state-of-the-art towards human-level efficiency.

For cyclic manipulation, Matsusaka et al. (2016) used resonance-based control to find the energy optimal constant stiffness for a pick-and-place task. Time-varying optimal stiffness modulation can be solved by optimal control approaches, as demonstrated in Nakanishi et al. (2011); Velasco et al. (2015). However, the

solvable optimal control formulation requires that the periodicity of movements holds perfectly. For a specific problem, *e.g.,* ball dribbling, Haddadin et al. (2018) developed a controller to dribble a ball stably using minimal peak power based on analysis of the stability of dribbling limit cycles and the effects of hand stiffness for robustness and energy efficiency.

Explosive Movement

An optimal control approach has also been used to exploit the energy storage effect for explosive movements. For instance, Braun et al. (2012) showed that discrete movements like throwing can be optimized by gradually feeding energy into the elastic elements before releasing it explosively for the throw, thereby amplifying the instantaneous power output, beyond what would otherwise be possible with the motors. Braun utilized a cost function to evaluate the distance of the ball being thrown. Such kind of problems can also be formulated as maximizing the output link velocity at the final time (Ozparpucu and Haddadin (2013)). Time-varying optimal stiffness profiles can be solved as a result of optimization to exploit the power amplification effect as discussed in §2.3. Energy efficiency demand in this case can be viewed as the need to output maximum instantaneous power using limited energy inflow.

More complex tasks have rarely been treated by an optimal control approach. One such example is the brachiation robot developed by Nakanishi et al. (2016), where the brachiation locomotion is formulated as a *multiphase* problem involving contacts and switching of dynamics. By optimizing the whole sequence the variable stiffness is ensured to be fully exploited.

However, this model-based method is difficult to generalize to a broad class of movements, because modelling of all kinds of complex movement skills is nearly impossible. A more plausible way is to generate the motion sequence given some movement primitives, (Hogan and Sternad (2012)). As argued by Hogan, both periodic movement and simple discrete movement can be viewed as two basic movement primitives. More complex movements can be produced by sequencing these basic ingredients. How to integrate optimal control with a movement primitives based framework is still an open question. *Sequential movement optimization*

in terms of variable impedance exploitation is an unexplored topic.

2.3.2 Reducing energy cost of stiffness modulation

Indeed, there is no unique optimal design of a VIA for different use-cases as they prioritize different functional features, such as maximum output power, stiffness range, accuracy, efficiency, *etc.*. For a general purpose VSA, the benefits of energy storage can only be enjoyed if there is efficient power flow between the motors and the elastic element (see Figure 2.5). However, in the early development of VSAs it was observed that the adjustment of stiffness causes high energy consumption. Even when $P_{\text{in2}} = 0$, the motor may still be consuming energy to maintain the elongation or compression of the elastic element. This has motivated theoretical studies of energy consumption for controlling stiffness *e.g.,* by Visser et al. (2010, 2011). Several energy-efficient designs of stiffness modulation mechanisms in VIAs were proposed recently. For instance, Jafari et al. (2013, 2014, 2015) and Groothuis et al. (2014) suggested a lever mechanism for adjusting stiffness and Braun et al. (2016, 2018) proposed a minimalistic stiffness modulator, both of which avoid having motor drives work against spring forces. Parallel springs are implemented in Plooij et al. (2016) and Jimenez-Fabian et al. (2017) to reduce required torque by locking potential energy into the parallel springs.

By minimizing the energy required for stiffness modulation, these actuators are particularly useful when there is a need of adjusting and sustaining stiffness at a constant position or load, *e.g.,* in protheses and exoskeleton (Braun et al. (2018)).

2.4 Energy Regeneration via Damping

2.4.1 Energy regeneration

The above examples addressed the energy efficiency in terms of *energy consumption* via either design or control.[6] An alternative is to focus on *energy regeneration*. From Figure 2.4, it can be seen that the link dissipates energy via

[6]When design and control are taken into account together, it can be called a co-design approach.

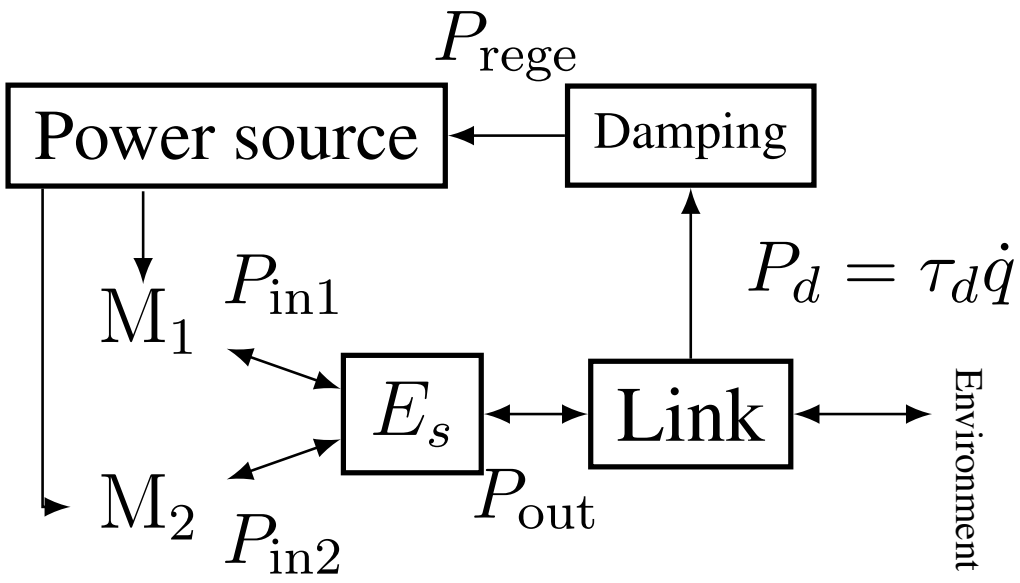


Figure 2.5: Schematic diagrams of the corresponding power flow of a VIA with regeneration through damping. The arrow between damping module and power source shows that the energy dissipated via damping can be harvested and used to recharge the power source.

damping elements, a uni-directional flow. If the latter primarily consist of frictional elements, this energy is wasted. However, if the damping mechanism is such that the dissipated energy can be harvested, this energy has the potential to be used to recharge the power source and decrease the overall net consumption. We illustrate the concept in Figure 2.5. So far, this possibility has received little attention in the field of compliant robotics.

Figure 2.6 depicts a conceptual diagram of a DC motor working as a generator. S1 and S2 are two controlled switches, When S1 switched on and S2 off, *regenerative braking* happens when motor M operates as generator and power flows into R, representing the power source as an electrical load. While when both S1, S2 switched on, no energy generated by M charges R and this is so-called *dynamical braking*. Another somehow dangerous case is *concurrent braking* when R also works as source and contributes to build up τ_m, resulting in large current and huge heat dissipation on the circuit. The switching frequency of S1 and S2 can be used to control the reflected damping property against mechanical load at motor's

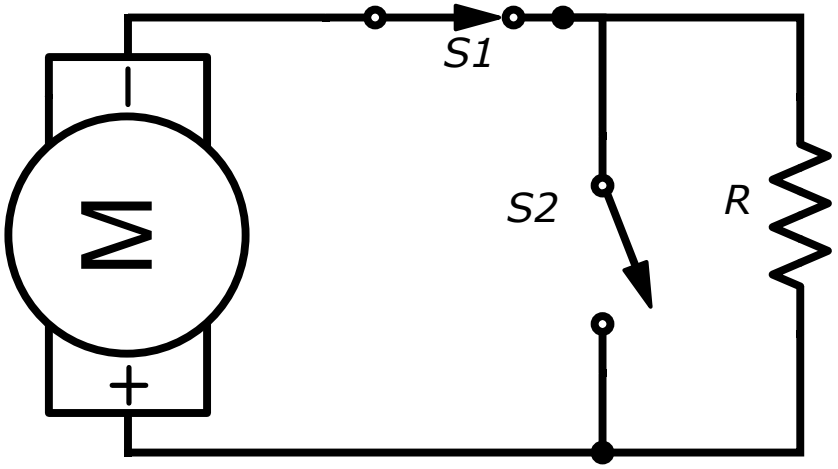


Figure 2.6: Diagram representing basic operation of DC motor damping as a generator. S1 and S2 are two controlled switches, R represents an electric load that works as power sink and store electrical energy. When S1 switched on and S2 off, *regenerative braking* happens when motor M operates as generator and power flow into R; while when both S1, S2 switched on, no energy generated by M charges R and this is so-called *dynamical braking*. Another somehow dangerous case is *concurrent braking* when R also works as source and contributes to build up τ_m, resulting in large current and huge heat dissipation on the circuit. The switching frequency of S1 and S2 can be used to control the reflected damping property against mechanical load at motor's output shaft.

output shaft.

A few applications in robotics can be found in the literature. For example, Seok et al. (2015) implemented regenerative electric motor drivers on the MIT Cheetah,[7] enabling the motors being used to both actuate and brake the joint, however, that actuation system is based on active impedance control, not physically compliant VIAs. The same regenerative braking principle is used as a kinetic energy harvester on a lower limb exoskeleton developed by Donelan et al. (2008), however, there the motor is used purely as a generator and plays no role as an actuator. Radulescu et al. (2012) showed the role that DC motor damping can play in generating braking force in a VIA, but did not explore its ability to harvest energy.

The fact that an electric motor can be used as a generator to convert kinetic energy into electricity and exert a braking effect was known as regenerative braking soon after the invention of electric railway in the late 19th century (Hellmund (1917)). This technology, which evolved into what is called Kinetic Energy Recovery System (KERS), has been widely used in modern vehicles driven particularly by electric motors (Gao et al. (1999)), or equipped with regenerative suspension systems (Zhang et al. (2018)).

2.4.2 Filling the Gap

However, the requirements for the regenerative damping system to be used with VIAs are different from the above use cases. First, for general purpose compliant actuators, the movement is generally bidirectional, but typical electrical energy storage elements (*e.g.,* batteries) are unidirectional, which indicates that a conversion mechanism is needed to ensure that the battery can be charged by energy from both directions of movement. This is different from the case where, in walking and running, the impact power from contact is basically unidirectional.

Secondly, regenerating energy in VIAs with variable physical damping couples the joint dynamics with the efficiency. Energy regeneration assigns an additional role for variable damping, thus more investigation is needed to determine appro-

[7]Note that, the Cheetah is not based on physical compliance but active impedance control, so that there is no elastic element between the link and motors that can store potential energy.

priate control strategies, in order to balance the trade-off between optimality of energy cost/regeneration and task achievement for specific tasks.

Admittedly, balancing the efficiency trade-off problem may increase the control complexity. There are some other drawbacks that should be taken into account when implementing regenerative damping. For instance, the hardware components may add more weight to the actuator and significantly influence the movement performance. Also, controlling energy harvesting requires more electronics components.

In the later chapters of the thesis it is shown how both issues can be addressed by (i) proposing a damping module design capable of harvesting energy from bi-directional movements, and (ii) evaluating optimal control as a means for dealing with the energy/task performance trade-off.

2.5 Summary

In this chapter, the basic concepts of VIAs were introduced. Comparison regarding energy efficiency between robots and living beings provides inspiration and insights into morphological intelligence embodiment. By looking at the recent development of compliant actuation, we can see an increasing number of researchers have been focusing on improving energy efficiency via design or control. Previous works in the literature have led to various types of novel stiffness modulation mechanisms that attempt to minimize the energy cost for adjusting and maintaining stiffness. Other designs can be found in the growing area of robotic locomotion and humanoid aim at efficient human-like locomotion. The associated control problems have been investigated for applications in walking robots, cyclic manipulation or explosive movements.

Another possible way to address the energy efficiency of VIAs, which has received little attention, is regeneration via damping. Especially when natural dynamics is hard to be exploited *e.g.,* for non-periodic discrete movements, a lot of energy would be dissipated and wasted by damping or friction. A plausible way to harvest this kind of energy is the use of regenerative braking, which has evolved into what today is called KERS in the industry of vehicles. In our view, both the elastic and regenerative damping could be kinetic energy harvesters. Intuitively,

regenerative damping can be viewed as a secondary energy recovery mechanism to trigger when springs can be fully exploited.

Furthermore, the optimal control approach for exploiting variable impedance has been limited to periodic or discrete movements. A broader class of sequential movements — which is a paradigm focusing on human-level manipulation — has not been considered in the field of VIAs. We have seen impressive designs of efficient legged robots that have emerged in the public eye. However, there is still much potential of energy efficiency improvement of general purpose VIAs, in particular for complex tasks involving upper limbs movements.

Chapter 3

Optimal Control Formulation

Throughout this thesis, optimal control is used as a mathematical tool to investigate the optimality and resolve the inherent actuation redundancy. This chapter introduces the modelling of system dynamics and optimal control framework. We also provide a closed form joint space tracking controller that automatically resolves actuation redundancy. The tools presented in this chapter will be used in later chapters.

3.1 Introduction

Optimal control is a widely used mathematical tool to naturally exploit variable impedance and resolve the actuation redundancy of compliant actuators (Garabini et al. (2011); Braun et al. (2012); Haddadin et al. (2012); Radulescu et al. (2012); Ozparpucu and Haddadin (2013, 2014); Zhakatayev et al. (2017)). It provides a rigorous framework to examine the optimality principle especially when an analytical solution is available. Efficient numerical solvers make it applicable to a broad class of problems commonly found in robot control.

3.2 System dynamics of VIAs

Throughout the thesis the type of compliant actuators of interest are VIAs whose mechanical stiffness and damping are both controllable. Previous works concerning the dynamics modelling of compliant robots are mostly based on the model presented by Spong (1987). The model was later used to derive control framework for elastic joint robots, *e.g.,* by Tomei (1991) and Albu-Schäffer et al. (2007). Although the analyses there were limited to constant elasticity, so that the control frameworks derived by them cannot be directly applied to VIAs, the basic model structure is the same that contains (a) link-side and (b) motor-side dynamics. In this chapter, we assume that the discussions are based on VIAs that have compliance in a series configuration as illustrated in Figure 2.3.

For a n degree of freedom (DOF) rigid-link robots, the joint configuration vector is defined as $\mathbf{q} \in \mathbb{R}^n$, $\mathbf{M} \in \mathbb{R}^{n \times n}$ denotes the inertia matrix which is positive-definite and symmetric, $\mathbf{G}$ is the gravitational vector, $\mathbf{C}$ is the Coriolis and centrifugal matrix. The link-side dynamics driven by compliant actuators can be represented by

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) = \boldsymbol{\tau}_a + \boldsymbol{\tau}_{\text{ext}}, \tag{3.1}$$

where $\boldsymbol{\tau}_a$ is the torque applied by the variable impedance actuators on the joints and $\boldsymbol{\tau}_{\text{ext}}$ represents the external torque. Before we discuss the modelling of dynamics including the actuators, which can be represented by the mass-spring-

damper model depicted in Figure 2.3, there are some assumptions need to be clarified.

Assumption 1. *The damper in the actuation system is arranged between the link and the base, and the damping d is independently controllable.*

Assumption 2. *The link-side and motor-side inertia are decoupled. Angular kinetic energy of each motor is only due to its own spinning.*

Assumption 3. *The elastic components have constant stiffness factors (e.g., linear springs) and satisfy Hooke's law.*

Given Assumption 1, we can have

$$\boldsymbol{\tau_a} = \boldsymbol{\tau}_s(\mathbf{q}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) - \mathbf{D}(\boldsymbol{\theta}_3)\dot{\mathbf{q}}, \tag{3.2}$$

where $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3 \in \mathbb{R}^n$ are configuration vectors for the equilibrium point (EP) motors, stiffness motors and dampers. $\mathbf{D}$ denotes the damping matrix. The first term of $\boldsymbol{\tau}_a$ is the torque $\boldsymbol{\tau}_s$ applied by the spring, and the second term represents a damping force generated by the damper, which is proportional to the joint velocity and controllable damping.

According to Assumption 1 and 2, there is no inertia coupling between the link and motors, and the inertia of damper can be included in the link-side inertia. With these simplifications the kinetic energy of the system can be written as

$$\mathcal{K} = \underbrace{\frac{1}{2}\dot{\mathbf{q}}^{\mathsf{T}}\mathbf{M}(\mathbf{q})\dot{\mathbf{q}}}_{\mathcal{K}_q} + \underbrace{\frac{1}{2}\dot{\boldsymbol{\theta}}_1^{\mathsf{T}}\mathbf{B_1}\dot{\boldsymbol{\theta}}_1}_{\mathcal{K}_{\theta 1}} + \underbrace{\frac{1}{2}\dot{\boldsymbol{\theta}}_2^{\mathsf{T}}\mathbf{B}_2\dot{\boldsymbol{\theta}}_2}_{\mathcal{K}_{\theta 2}}, \tag{3.3}$$

where $\mathbf{B}_1, \mathbf{B}_2$ are constant and diagonal matrices that contains the inertia of the EP and stiffness motors, $\mathcal{K}_q, \mathcal{K}_{\theta 1}, \mathcal{K}_{\theta 2}$ represent the kinetic energy of the links, EP motors and stiffness motors respectively. The potential energy due to gravity is denoted by $\mathcal{U}_g$ and only depends on the joint configuration $\mathbf{q}$.

Without loss of generality, L_i and L_{i0} denote the length and equilibrium length of the elastic element in the i-th DOF, and s_i is the corresponding elasticity constant. According to Assumption 3, the restorative force of the i-th elastic element is $F_{s,i} = -s_i(L_i - L_{i0})$. The equilibrium length L_{i0} is a constant determined by

the design of the actuation mechanism, and the force $F_{s,i}$ is adjusted by the length L_i that depends on the joint position q_i and the configuration variables $(\theta_{1,i}, \theta_{2,i})$. We denote $h_i(q_i, \theta_{1,i}, \theta_{2,i}) = L_i - L_{i0}$ as the effective length of deformation of the i-th elastic element. The elastic potential energy can be represented by

$$\mathcal{U}_e(\mathbf{q}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \frac{1}{2}\mathbf{h}^\mathsf{T}\mathbf{S}\mathbf{h}, \tag{3.4}$$

where $\mathbf{S} = \mathrm{diag}(s_i) \in \mathbb{R}^n$, $\mathbf{h} = (h_1, ..., h_n)^\mathsf{T} \in \mathbb{R}^n$.

Then we can define a generalized coordinates vector $\boldsymbol{\rho} = (\mathbf{q}^\mathsf{T}, \boldsymbol{\theta}_1^\mathsf{T}, \boldsymbol{\theta}_2^\mathsf{T})^\mathsf{T} \in \mathbb{R}^{3n}$ and write the Lagrangian function of the system as

$$\mathcal{L}(\boldsymbol{\rho}) = \mathcal{K} - \mathcal{U}_e - \mathcal{U}_g. \tag{3.5}$$

By applying the Lagrangian dynamics formulation (Lynch and Park (2017), pp. 271-279) we have

$$\frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{\rho}}} - \frac{\partial \mathcal{L}}{\partial \boldsymbol{\rho}} = \mathcal{T}, \tag{3.6}$$

where $\mathcal{T} = (\boldsymbol{\tau}_{\text{ext}}^\mathsf{T}, \boldsymbol{\tau}_{m1}^\mathsf{T}, \boldsymbol{\tau}_{m2}^\mathsf{T})^\mathsf{T} \in \mathbb{R}^{3n}$ is the generalized torques. $\boldsymbol{\tau}_{m1}, \boldsymbol{\tau}_{m2}$ are torques applied on the EP and stiffness motors respectively. Note that, it is assumed that the vector derivatives are column vectors in this section for better readability of mathematical expressions.

By substituting (3.5), (3.6) becomes

$$\begin{pmatrix} \frac{\partial}{\partial t}\frac{\partial \mathcal{K}_q}{\partial \dot{\mathbf{q}}} \\ \frac{\partial}{\partial t}\frac{\partial \mathcal{K}_{\theta 1}}{\partial \dot{\boldsymbol{\theta}}_1} \\ \frac{\partial}{\partial t}\frac{\partial \mathcal{K}_{\theta 2}}{\partial \dot{\boldsymbol{\theta}}_2} \end{pmatrix} - \begin{pmatrix} \frac{\partial \mathcal{K}_q}{\partial \mathbf{q}} - \frac{\partial \mathcal{U}_e}{\partial \mathbf{q}} - \frac{\partial \mathcal{U}_g}{\partial \mathbf{q}} \\ -\frac{\partial \mathcal{U}_e}{\partial \boldsymbol{\theta}_1} \\ -\frac{\partial \mathcal{U}_e}{\partial \boldsymbol{\theta}_2} \end{pmatrix} = \mathcal{T}. \tag{3.7}$$

Since $\mathbf{M}$ is a symmetric positive-definite matrix, $\partial \mathcal{K}_q / \partial \dot{\mathbf{q}} = \mathbf{M}(\mathbf{q})\dot{\mathbf{q}}$. $\mathcal{K}_{\theta 1}$ and $\mathcal{K}_{\theta 2}$ are independent of $\mathbf{q}$, thus $\frac{\partial}{\partial t}\frac{\partial \mathcal{K}_{\theta i}}{\partial \boldsymbol{\theta}_i} = \mathbf{B}_i\ddot{\boldsymbol{\theta}}_i, i = 1, 2$. One can obtain the equation

of motion by applying the chain rule and rearranging (3.7):

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) = \boldsymbol{\tau}_s + \boldsymbol{\tau}_{\text{ext}} \tag{3.8}$$

$$\mathbf{B}_1\ddot{\boldsymbol{\theta}}_1 = \boldsymbol{\tau}_{l1} + \boldsymbol{\tau}_{m1} \tag{3.9}$$

$$\mathbf{B}_2\ddot{\boldsymbol{\theta}}_2 = \boldsymbol{\tau}_{l2} + \boldsymbol{\tau}_{m2} \tag{3.10}$$

where $\boldsymbol{\tau}_s$ are torques applied by springs on joints, $\boldsymbol{\tau}_{l1,2}$ represent the motor load torques exerted by the springs, and

$$\boldsymbol{\tau}_s = -\frac{\partial \mathcal{U}_e}{\partial \mathbf{q}} = -\frac{\partial \mathbf{h}}{\partial \mathbf{q}}\mathbf{Sh} \tag{3.11}$$

$$\boldsymbol{\tau}_{l1} = -\frac{\partial \mathcal{U}_e}{\partial \boldsymbol{\theta}_1} = -\frac{\partial \mathbf{h}}{\partial \boldsymbol{\theta}_1}\mathbf{Sh} \tag{3.12}$$

$$\boldsymbol{\tau}_{l2} = -\frac{\partial \mathcal{U}_e}{\partial \boldsymbol{\theta}_2} = -\frac{\partial \mathbf{h}}{\partial \boldsymbol{\theta}_2}\mathbf{Sh} \tag{3.13}$$

$$\mathbf{G}(\mathbf{q}) = \frac{\partial \mathcal{U}_g}{\partial \mathbf{q}} \tag{3.14}$$

$$\mathbf{C}\dot{\mathbf{q}} = \frac{\partial \mathbf{M}(\mathbf{q})}{\partial t}\dot{\mathbf{q}} - \frac{\partial \mathcal{K}_q}{\partial \mathbf{q}} \tag{3.15}$$

In addition, dissipative terms can be added to the equation, which are the damping force provided by the actuators, and the frictional forces. For simplicity, we assume that the link-side frictions are included in matrix $\mathbf{C}$. Finally, the equation of motion can be written as

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) = \boldsymbol{\tau}_a + \boldsymbol{\tau}_{\text{ext}} \tag{3.16}$$

$$\mathbf{B}_1\ddot{\boldsymbol{\theta}}_1 + \mathbf{B}_{f1}\dot{\boldsymbol{\theta}}_1 = \boldsymbol{\tau}_{l1} + \boldsymbol{\tau}_{m1} \tag{3.17}$$

$$\mathbf{B}_2\ddot{\boldsymbol{\theta}}_2 + \mathbf{B}_{f2}\dot{\boldsymbol{\theta}}_2 = \boldsymbol{\tau}_{l2} + \boldsymbol{\tau}_{m2} \tag{3.18}$$

To investigate the stiffness characteristic of the actuators, the stiffness matrix $\mathbf{K} \in \mathbb{R}^{n\times n}$ is defined as

$$\mathbf{K}(\mathbf{q}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = -\frac{\partial \boldsymbol{\tau}_s}{\partial \mathbf{q}}. \tag{3.19}$$

3.2.1 Simplification of motor dynamics

In practice, the motor dynamics can be simplified if the motors are position-controlled and have high ratio gearboxes. A 2nd-order critically-damped dynamics is introduced in Braun et al. (2013) :

$$\ddot{\boldsymbol{\theta}} = \boldsymbol{\beta}^2(\mathbf{u} - \boldsymbol{\theta}) - 2\boldsymbol{\beta}\dot{\boldsymbol{\theta}} \tag{3.20}$$

Usually, the coefficient β is chosen empirically to fit the step response of the servomotors. In effect, it imposes a constraint on the acceleration specified by β as a bandwidth limit. Also, it naturally imposes a position constraint on θ by constraining $\mathbf{u}$. The benefit is that the control constraint is easy to be incorporated in a numerical algorithm *e.g.,* as shown by Tassa et al. (2014). Actually, the simplification presumes the motor is controlled by a D-P type controller such that the motor dynamics can be reduced to 2nd order (Appendix B). Another assumption is that the load torque applies no work on the motor, in other word, the motor only outputs but receive no power back. This is true for non-backdrivable servomotors.

3.3 Optimal control formulation

To formulate the optimal control problem the robot dynamics is re-written into the form of state-space model (SSM)

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}), \tag{3.21}$$

where $\mathbf{x} = (\mathbf{q}^\mathsf{T}, \dot{\mathbf{q}}^\mathsf{T}, \boldsymbol{\theta}_1^\mathsf{T}, \dot{\boldsymbol{\theta}}_1^\mathsf{T}, \boldsymbol{\theta}_2^\mathsf{T}, \dot{\boldsymbol{\theta}}_2^\mathsf{T})^\mathsf{T} \in \mathbb{R}^{6n}$ denotes the state vector, $\mathbf{u} \in U \in \mathbb{R}^m$ is the control input. $\mathbf{u}$ is constrained by the set of admissible controls denoted as:

$$U = \{\mathbf{u} \in \mathbb{R}^{\mathrm{m}} : \mathbf{u}_{\min} \preceq \mathbf{u} \preceq \mathbf{u}_{\max}\}, \tag{3.22}$$

where $\mathbf{u}_{\min}$ and $\mathbf{u}_{\max}$ are the lower and upper bounds on $\mathbf{u}$, representing the physical constraints of the robot control inputs. Moreover, $\mathbf{u} = (\mathbf{u}_1^\mathsf{T}, \mathbf{u}_2^\mathsf{T}, \mathbf{u}_3^\mathsf{T})^\mathsf{T}$ contains the control inputs for the EP motors, stiffness motors, and damping. $\mathbf{u}_1, \mathbf{u}_2$ have different meaning depending on if the motors are position, velocity or torque con-

trolled. For the above simplified motor dynamics, $\mathbf{u}_1, \mathbf{u}_2$ represents the position command.

In general, the finite horizon optimal control problem is to find an optimal $\mathbf{u} = \mathbf{u}(t, \mathbf{x}) \in U$ that minimizes a cost functional J within a finite time interval $t \in [0, t_f]$. Mathematically it can be written as:

$$\begin{aligned} \text{Minimize} \quad & J \quad \text{over } \{\mathbf{u}(t)\}_{t\in[0,t_f)} \\ \text{subject to} \quad & \mathbf{u}(t) \in U \\ & \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \\ & \mathbf{x}(0) = \mathbf{x}_0 \end{aligned} \tag{3.23}$$

J is usually in the form of

$$J = H(\mathbf{x}(t_f), t_f) + \int_0^{t_f} l(\mathbf{x}(t), \mathbf{u}(t), t)\,\mathrm{d}t \tag{3.24}$$

where $H(\mathbf{x}(t_f), t_f)$ is the final cost at terminal time t_f and l represents the immediate cost.

3.3.1 Objective functions

The immediate cost $l(t)$ can be divided into a task term $l_t(t)$ and an energy term $l_e(t)$. To generate desired behaviour fulfilling a specific task requirement, one can assign a cost function quantifying the task performance. The definition of task term is of course task-specific. In many problems the primary task objective is only defined by the terminal cost H. A common example is reaching a target:

$$H(\mathbf{x}(t_f), t_f) = (\mathbf{x}^{\mathsf{T}}(t_f) - \mathbf{x}^*)\,\mathbf{H}_f\,(\mathbf{x}(t_f) - \mathbf{x}^*) \tag{3.25}$$

where $\mathbf{H}_f$ is a diagonal matrix and $\mathbf{x}^*$ is the target state. In such case $l_t(t) = 0$.

The mostly used energy term is a quadratic control effort

$$l_e(\mathbf{u}) = \mathbf{u}^T \mathbf{H}_e \mathbf{u}. \tag{3.26}$$

In robotics, $\mathbf{u}$ is often the joint torque for many torque-controlled robots, and the

control effort is simply squared torques. For biomechanical models $\mathbf{u}$ is usually the muscle forces. In both cases this is a meaningful energy cost and is why it is called effort. However, for VIAs, as introduced above, $\mathbf{u}$ can be a position command and thus the effort here makes no sense. However, a simple variant can make it still useful:

$$l_e(\mathbf{u}) = (\mathbf{u}_1 - \mathbf{q}^*) + \mathbf{u}_2 + \mathbf{u}_3 \tag{3.27}$$

which means the minimal energy value for EP motors should be the target joint position.

It is well-known that designing appropriate cost functions is not trivial (Uno et al. (1989); Todorov (2004)). In terms of seeking energy-optimal solutions, how to calculate the energy cost and balance the trade-off with primary task goal requires a lot of domain-specific knowledge. For example, in many explosive athletic movements the cost objective is fully defined by kinematic attributes and there is no need to minimize energy cost (Pandy et al. (1990)). It also has been demonstrated in Braun et al. (2012) by throwing experiment that the maximal throwing distance of a 2-link Mechanically Adjustable Compliance and Controllable Equilibrium Position Actuator (MACCEPA) arm is achieved when control effort term vanishes.

In many scenarios, energy consumption is of concern. There are studies in both biomechanics Nelson (1983) and robotics Remy et al. (2012) examining the effects of different energetic costs on emergent behaviours. Even if the terms to be included in the cost function are determined, it still remains a question how to tune the weighting for different terms, especially when they are competing with each other.

To investigate the use of realistic energy measurement such as mechanical and electrical work, we conducted a preliminary study based on simulation using various energetic costs, which is presented in Appendix C. The result there indicates that by using control effort, without the information of real energy consumption, it fails to reach the real limit, *i.e.,* the efficient frontier of energy efficiency. However, by tuning the cost function weighting, with simple control effort we can still adjust the balance of performance-cost trade-off. Also, in practice the quadratic control effort makes the optimization more robust and results in smooth traject-

ory. In chapter 6, we will discuss the possibility of using simple control effort for low-level control and realistic energy measure for high-level planning. By doing so, it leverages the benefits of both.

3.3.2 Example: MACCEPA-VD

MACCEPA proposed by Van Ham et al. (2007) modulates the stiffness by adjusting the spring pretension. It was implemented for various applications, *e.g.,* a bipedal walking robot in VanderBorght et al. (2008) and explosive movement experiments by Braun et al. (2012). Radulescu et al. (2012) incorporated a variable damping mechanism based on eddy current principle into MACCEPA. In the MACCEPA with variable damping (MACCEPA-VD), the equilibrium position and joint stiffness are controlled by two servomotors separately, while the damping coefficient is modulated by changing the duty cycle of the circuit for braking a DC motor attached rigidly to the joint. The system is illustrated in Figure 3.1. The forward dynamics for this single joint system can be written as:

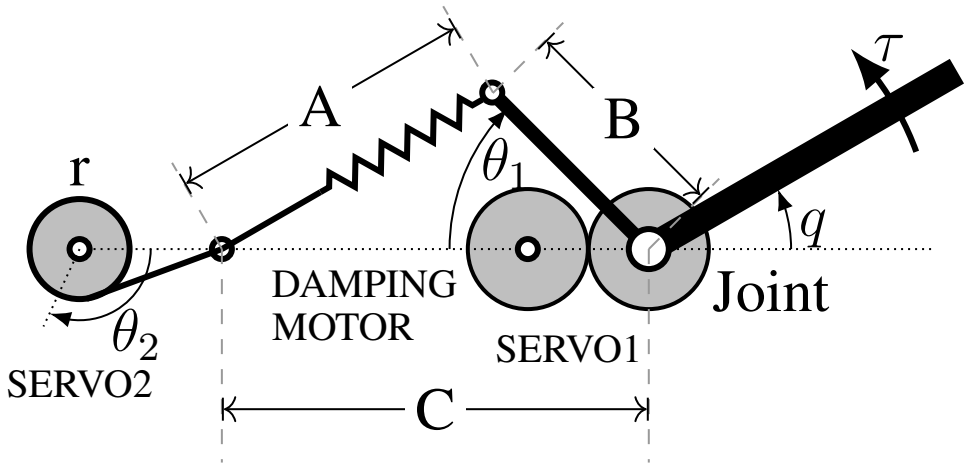


Figure 3.1: Diagram of MACCEPA Van Ham et al. (2007) with variable damping Radulescu et al. (2012).

$$\ddot{q} = (\tau_s - d(u_3)\dot{q} - b\dot{q} - \tau_{\text{ext}})m^{-1} \tag{3.28}$$

$$\ddot{\theta}_1 = \beta^2(u_1 - \theta_1) - 2\beta\dot{\theta}_1 \tag{3.29}$$

$$\ddot{\theta}_2 = \beta^2(u_2 - \theta_2) - 2\beta\dot{\theta}_2 \tag{3.30}$$

where $q, \dot{q}, \ddot{q}$ are the joint angle, velocity and acceleration, respectively, b is the viscous friction coefficient for the joint, m is the link inertia, τ_s is the torque generated by the spring force, and τ_{ext} is the joint torque due to external loading (the following reports results for the case of no external loading, *i.e.,* $\tau_{\text{ext}} = 0$). $\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2, \ddot{\theta}_1, \ddot{\theta}_2$ are the motor angles, velocities and accelerations.

The motor angles θ_1, θ_2 and damping d are controlled by control input $\mathbf{u} = (u_1, u_2, u_3)^{\mathsf{T}}$. The servomotor dynamics (3.29), (3.30) are assumed to behave as a critically damped system, with β constraining the maximum acceleration of the 2nd order dynamical system.

The spring length L and the rest length L_0 are given by Van Ham et al. (2007) as:

$$L = A(q, \theta_1) + r\theta_2, \tag{3.31}$$

$$L_0 = C - B, \tag{3.32}$$

where $A(q, \theta_1) = \sqrt{B^2 + C^2 - 2BC\cos(\theta_1 - q)}$, B and C are the lengths shown in Figure 3.1, r is the radius of the winding drum used to adjust the spring pre-tension. The linear spring constant is denoted by κ.

The torques $\tau_s, \tau_{l_1}, \tau_{l_2}$ can be calculated as follows by using (3.11) - (3.13):

$$\tau_s = \kappa BC\sin(\theta_1 - q)(1 + \frac{r\theta_2 - |C - B|}{A(q, \theta_1)}) \tag{3.33}$$

$$\tau_{l_1} = \tau_s \tag{3.34}$$

$$\tau_{l_2} = \kappa(r\theta_2 - |C - B| + A(q, \theta_1)) \tag{3.35}$$

The stiffness of this system depends on the joint and motor positions q, θ_1, θ_2. By applying (3.19) on τ_s it yields

$$k(q, \theta_1, \theta_2) = \kappa BC\cos(\theta_1 - q)(1 + \frac{r\theta_2 - |C - B|}{A}) - \frac{\kappa B^2C^2\sin^2(\theta_1 - q(r\theta_2 - |C - B|))}{A^{\frac{3}{2}}}. \tag{3.36}$$

It can be seen that the same spring pretension θ_2 can result in different joint stiffness under different joint and EP configurations.

The damping coefficient $d(u_3)$ depends on control input u_3 and the functional form is determined by the damping mechanism. The simplest case is a linear damping function:

$$d(u_3) = \bar{d}u_3, \tag{3.37}$$

where $\bar{d}$ is maximum damping coefficient and the control input varies from 0 to 1 ($u_3 \in [0, 1]$). More detailed calculation of the damping effect provided by a DC motor is covered in Appendix B.

The corresponding SSM of MACCEPA-VD can be written as

$$\mathbf{f} = \begin{cases} x_2 \\ (\tau_s(x_1, x_2, x_3) - (d(u_3) + b)x_2)m^{-1} \\ x_5 \\ x_6 \\ \beta^2(u_1 - x_3) - 2\beta x_5 \\ \beta^2(u_2 - x_4) - 2\beta x_6 \end{cases} \tag{3.38}$$

where $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6)^\top = (q, \dot{q}, \theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2)^\top \in \mathbb{R}^6$ is the state vector, $\mathbf{u} = (u_1, u_2, u_3)^\top \in \mathbb{R}^3$ is the control input.

3.4 Iterative Linear Quadratic Regulator

One way to solve optimal control problem is to derive the Hamilton-Jacobi-Bellman (HJB) equation starting with bellman equation (Bellman (1957)). The HJB equation provides sufficient condition for the global optimal solution however it suffers from the "curse of dimensionality" because it requires discretization of state-space, by which the number of discrete states grows exponentially in the state-space dimensionality.

Another method, which only works for deterministic dynamics, is the Pontryagin's Maximum Principle (PMP) Kirk (1970). By PMP the problem is converted to find a solution of a non-linear boundary value problem. The ODE can then be solved by numerical optimization via gradient descent.

A more efficient solver is provided by Dynamic Differential Programming

(DDP), more specifically its variant Iterative Linear Quadratic Regulator (ILQR) (Li and Todorov (2004)) that use first order approximation for dynamics. Tassa et al. (2014) recently extended it by incorporating inequality constraints on controls and solving a quadratic programming problem. The ILQR method has been demonstrated to be more computationally efficient than DDP and PMP (solving ODE with gradient descent), although the solutions are also only locally optimal; and it is applicable to problems with non-quadratic cost function.

The ILQR method is initialized with a nominal control sequence $\hat{\mathbf{u}}$ and corresponding trajectory $\hat{\mathbf{x}}$. These are then iteratively improved by solving a set of local LQR sub-problems. The sub-problems are formed by linear approximation of the dynamics and quadratic approximation of the objective function,

$$\delta\dot{\mathbf{x}} = \mathbf{f}_x\delta\mathbf{x} + \mathbf{f}_u\delta\mathbf{u}, \tag{3.39}$$

$$\begin{aligned}\Delta J = H_{\mathbf{x}}^{\mathsf{T}}\delta\mathbf{x}(T) + \delta\mathbf{x}^{\mathsf{T}}H_{\mathbf{xx}}\delta\mathbf{x}(T) + \int_0^T (l_{\mathbf{x}}^{\mathsf{T}}\delta\mathbf{x} + l_{\mathbf{u}}^{\mathsf{T}}\delta\mathbf{u})dt \\ + \frac{1}{2}\int_0^T (\delta\mathbf{x}^{\mathsf{T}}l_{\mathbf{xx}}\delta\mathbf{x} + \delta\mathbf{x}^{\mathsf{T}}l_{\mathbf{xu}}\delta u + \delta\mathbf{u}^{\mathsf{T}}l_{\mathbf{uu}}\delta\mathbf{u})dt,\end{aligned} \tag{3.40}$$

$(\delta\mathbf{x}, \delta\mathbf{u})$ is obtained by solving the sub-problem using a modified Ricatti-like system Li and Todorov (2007). $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ can be updated by $\hat{\mathbf{x}} \leftarrow \hat{\mathbf{x}} + \delta\mathbf{x}$ and $\hat{\mathbf{u}} \leftarrow \hat{\mathbf{u}} + \delta\mathbf{u}$. When the iterations converge, the optimal solution $(\mathbf{x}^*, \mathbf{u}^*)$ is returned with feedback gains $\mathbf{L}^*(t)$. The feedback control law for optimal task execution is defined as

$$\mathbf{u}_{opt}(t, \mathbf{x}) = \mathbf{u}^*(t) + \mathbf{L}^*(t)(\mathbf{x} - \mathbf{x}^*(t)). \tag{3.41}$$

In our implementation, we compute the analytical derivatives of quadratic control terms. For other non-quadratic cost functions and robot dynamics, we utilize a finite difference approximation. The forward pass of dynamics is computed with the Runge-Kutta Fourth (RK4) method. As the solution given by DDP is locally optimal, it needs multiple runs with random initialization to find the global optimum. However, sometimes domain-specific knowledge helps with designing a "good" initialization. For instance, to optimize a reaching movement, we can

generally let $\hat{\mathbf{u}}_1^{\text{init}} = \mathbf{q}^*$ to guide the EP motor to move towards the target joint angles.

3.5 Extended Inverse Dynamics Controller

For a manipulator with redundant degrees of freedom the inverse Jacobian of kinematics does not exist. Therefore, inverse kinematics control usually is realized by using the pseudo inverse technique to track the trajectory in the workspace. A popular example is the one suggested by Khatib (1987). The null-space term in the inverse kinematic controller allows considering the order of task priority (Nakamura et al. (1987)). In addition, the inverse dynamics gives the mapping from joint space to controlled torques. Traditional robots have one torque source for each joint so that there is no redundancy issue in this joint to torque mapping. For over-actuated manipulators where the torques are linear combination of "muscle forces", *e.g.,* $\boldsymbol{\tau} = \mathbf{W}\mathbf{u}$, by applying the pseudo inverse $\mathbf{W}^\dagger$ the controller can be easily extended to "muscle space".

However, from the MACCEPA-VD we presented in this chapter, the torques generated by many VIAs are non-linear in control. Here an extension of inverse dynamics solutions is presented to enable joint trajectory tracking with VIAs. Tracking EP and desired stiffness has been addressed by Howard et al. (2011). Solutions for elastic joint robots with constant stiffness was given by Albu-Schäffer et al. (2007). To the author's knowledge, this closed form joint space tracking controller has not been introduced in the literature for VIAs.

3.5.1 Resolve actuation redundancy

Joint trajectory tracking with torque feedback

Assume that a reference joint trajectory has been given for tracking control. With the information of desired position, velocity and acceleration $\{\mathbf{q}_{\text{des}}, \dot{\mathbf{q}}_{\text{des}}, \ddot{\mathbf{q}}_{\text{des}}\}$ and the robot dynamics, we can use inverse dynamics to calculate the desired joint torque $\boldsymbol{\tau}_{\text{des}}$ at each time step. Within this section, the torque applied by the actuator is denoted as $\boldsymbol{\tau}_a(\mathbf{q}, \dot{\mathbf{q}}, \boldsymbol{\theta})$, which is a function depending on the joint state

and internal control variables. The tracking task is then converted to be in the form of a constraint description: $\boldsymbol{\tau}_a - \boldsymbol{\tau}_{\text{des}} = \mathbf{0}$. In order to make the left-hand side converge to zero asymptotically stably, we assume that it fulfils

$$(\dot{\boldsymbol{\tau}}_a - \dot{\boldsymbol{\tau}}_{\text{des}}) + \mathbf{K}_p(\boldsymbol{\tau}_a - \boldsymbol{\tau}_{\text{des}}) = \mathbf{0} \tag{3.42}$$

where $\mathbf{K}_p$ is a proportional gain matrix.

Taking the derivatives of the torque function yields

$$\begin{aligned}\dot{\boldsymbol{\tau}}_a &= \frac{\partial \boldsymbol{\tau}_a}{\partial \boldsymbol{\theta}}\dot{\boldsymbol{\theta}} + \frac{\partial \boldsymbol{\tau}_a}{\partial \mathbf{q}}\dot{\mathbf{q}} + \frac{\partial \boldsymbol{\tau}_a}{\partial \dot{\mathbf{q}}}\ddot{\mathbf{q}} \\ &= \mathbf{J}_{\boldsymbol{\theta}}\dot{\boldsymbol{\theta}} + \mathbf{J}_{\mathbf{q}}\dot{\mathbf{q}} + \mathbf{J}_{\dot{\mathbf{q}}}\ddot{\mathbf{q}}\end{aligned} \tag{3.43}$$

where $\mathbf{J}_{\boldsymbol{\theta}}, \mathbf{J}_{\mathbf{q}}, \mathbf{J}_{\dot{\mathbf{q}}}$ are used to represent the Jacobians of $\boldsymbol{\tau}_a$ w.r.t. $\boldsymbol{\theta}, \mathbf{q}, \dot{\mathbf{q}}$. For the following derivation we assume that $\boldsymbol{\theta}$ is controlled in velocity domain, *i.e.,* $\mathbf{u} = \dot{\boldsymbol{\theta}}$ is the control vector.

By rearranging (3.42) and substituting (3.43) we can have

$$\mathbf{J}_{\boldsymbol{\theta}}\mathbf{u} = \dot{\boldsymbol{\tau}}_{\text{des}} - \mathbf{J}_{\mathbf{q}}\dot{\mathbf{q}} - \mathbf{J}_{\dot{\mathbf{q}}}\ddot{\mathbf{q}} + \mathbf{K}_p(\boldsymbol{\tau}_{\text{des}} - \boldsymbol{\tau}_a) \tag{3.44}$$

which is in form of $\mathbf{A}\mathbf{u} = \mathbf{b}$, where

$$\mathbf{A} = \mathbf{J}_{\boldsymbol{\theta}}, \tag{3.45}$$

$$\mathbf{b} = \dot{\boldsymbol{\tau}}_{\text{des}} - \mathbf{J}_{\mathbf{q}}\dot{\mathbf{q}} - \mathbf{J}_{\dot{\mathbf{q}}}\ddot{\mathbf{q}} + \mathbf{K}_p(\boldsymbol{\tau}_{\text{des}} - \boldsymbol{\tau}_a). \tag{3.46}$$

Given a control cost metric matrix N, by applying Lemma A.4 the control law can be given as

$$\mathbf{u} = \mathbf{N}^{-\frac{1}{2}}(\mathbf{J}\mathbf{N}^{-\frac{1}{2}})^{\dagger}\mathbf{b} + \mathbf{N}^{-\frac{1}{2}}(\mathbf{I} - (\mathbf{J}\mathbf{N}^{-\frac{1}{2}})^{\dagger}\mathbf{J}\mathbf{N}^{-\frac{1}{2}})\mathbf{N}^{\frac{1}{2}}\mathbf{u}_1 \tag{3.47}$$

where the subscript of $\mathbf{J}_{\boldsymbol{\theta}}$ is omitted for readability.

In this controller, the feedback term computes the tracking error of torque, which requires torque sensor feedback. However, thanks to the fact that the spring torque of VSAs can be estimated by states of joint and motors, in hardware the implementation of a torque sensor is not necessary. An additional encoder for

measuring the spring length can enhance the accuracy of torque estimation.

Joint trajectory tracking with joint feedback

In the controller described by (3.47), the desired joint trajectory is transformed to desired torque. The drawback is that it includes only a positional feedback term which may be hard to tune a suitable gain to ensure stability for all trajectories. A controller that explicitly include joint state feedback may be preferred if tracking accuracy and stability is critical.

Suppose that the robot is asked to track a desired trajectory and satisfies a stability condition

$$(\ddot{\mathbf{q}} - \ddot{\mathbf{q}}_{\text{des}}) + \mathbf{K}_3(\ddot{\mathbf{q}} - \ddot{\mathbf{q}}_{\text{des}}) + \mathbf{K}_2(\dot{\mathbf{q}} - \dot{\mathbf{q}}_{\text{des}}) + \mathbf{K}_1(\mathbf{q} - \mathbf{q}_{\text{des}}) = \mathbf{0} \tag{3.48}$$

Then taking derivatives of

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) = \boldsymbol{\tau}_a(\mathbf{q}, \boldsymbol{\theta}) \tag{3.49}$$

yields

$$\mathbf{M}\dddot{\mathbf{q}} + \frac{\partial \mathbf{M}}{\partial t}\ddot{\mathbf{q}} + \mathbf{C}\ddot{\mathbf{q}} + \frac{\partial \mathbf{C}}{\partial t}\dot{\mathbf{q}} + \frac{\partial \mathbf{G}}{\partial t} = \mathbf{J}_{\boldsymbol{\theta}}\dot{\boldsymbol{\theta}} + \mathbf{J}_{\mathbf{q}}\dot{\mathbf{q}} + \mathbf{J}_{\dot{\mathbf{q}}}\ddot{\mathbf{q}}. \tag{3.50}$$

Combining (3.48) and (3.50), after rearranging them, we get

$$\begin{aligned}\mathbf{J}_{\boldsymbol{\theta}}\mathbf{u} = \mathbf{M}\dddot{\mathbf{q}}_{\text{des}} - \mathbf{J}_{\mathbf{q}}\dot{\mathbf{q}} - \mathbf{J}_{\dot{\mathbf{q}}}\ddot{\mathbf{q}} + \frac{\partial \mathbf{M}}{\partial t}\ddot{\mathbf{q}} + \mathbf{C}\ddot{\mathbf{q}} + \frac{\partial \mathbf{C}}{\partial t}\dot{\mathbf{q}} + \frac{\partial \mathbf{G}}{\partial t} \\ + \mathbf{M}(\mathbf{K}_3(\ddot{\mathbf{q}} - \ddot{\mathbf{q}}_{\text{des}}) + \mathbf{K}_2(\dot{\mathbf{q}} - \dot{\mathbf{q}}_{\text{des}}) + \mathbf{K}_1(\mathbf{q} - \mathbf{q}_{\text{des}}))\end{aligned} \tag{3.51}$$

Still, the right-hand side is $\mathbf{b}$, the controller is given by (3.47). In practice, it only needs joint state feedback.

In this section, two controllers are derived to show how the actuation redundancy can be resolved to track a trajectory in joint space. The former one includes toque feedback, whereas the latter uses joint position feedback. In practice, the actuator torque and joint position feedback may both be available. The selection of the controller depends on applications. Below is example of a point-to-point

reaching task which compares these two controllers in terms of tracking a minimal jerk trajectory and a trajectory generated by optimal control.

3.5.2 Examples

A point-to-point reaching task of MACCEPA-VD is used as example. Two reference joint trajectories are generated by analytical minimal jerk formula (Flash and Hogan (1985)) and optimal control. The controller with torque feedback is used to track minimal-jerk trajectory with gain $K_p = 10$, and the optimal control trajectory with gain $K_p = 50$. Both gains are tuned manually for best performance. It shows the drawback that for this controller the performance is very sensitive to the gain. It can be seen that the desired acceleration of OC trajectory is not as smooth as that of the minimal-jerk trajectory. As a consequence, it requires larger gain to ensure the tracking accuracy.

Next we implement the tracking controller with joint state feedback with gain $K_1, K_2, K_3 = 1000, 500, 50$. As shown in Figure 3.3 it demonstrates good tracking performance for both minimal-jerk trajectory and OC trajectory. The joint tracking stability is ensured even if the velocity and acceleration profiles when tracking the OC trajectory can not match the reference perfectly. The velocity and acceleration tracking shows a significant phase delay because of the model mismatch. In our optimal control model the acceleration limit is constrained by the coefficient β of the 2nd order motor dynamics. While in the controller we explicitly add a hard constraint of $10\,\mathrm{rad/s}$. The resulting trajectory thus cannot break the velocity limit when tracking the reference. Nevertheless, the joint state feedback terms involved in the controller guarantee the stability and good tracking performance approaching the target position.

In addition, the null-space controller bu_1 is exploited in the example shown in Figure 3.3. We define $\mathbf{u}_1 = (-(\theta_1 - q^*), -\theta_2, 0)^\mathsf{T}$ to drive the motors towards desired positions.[1] This null-space term will naturally guide EP towards the target joint angle, and relax the spring by driving the stiffness motor back to the zero position. This simple example shows the effectiveness of the extended inverse dynamics controller and the potential for null-space exploitation. This opens the

[1] Note that $\mathbf{u}_1$ is in velocity domain.

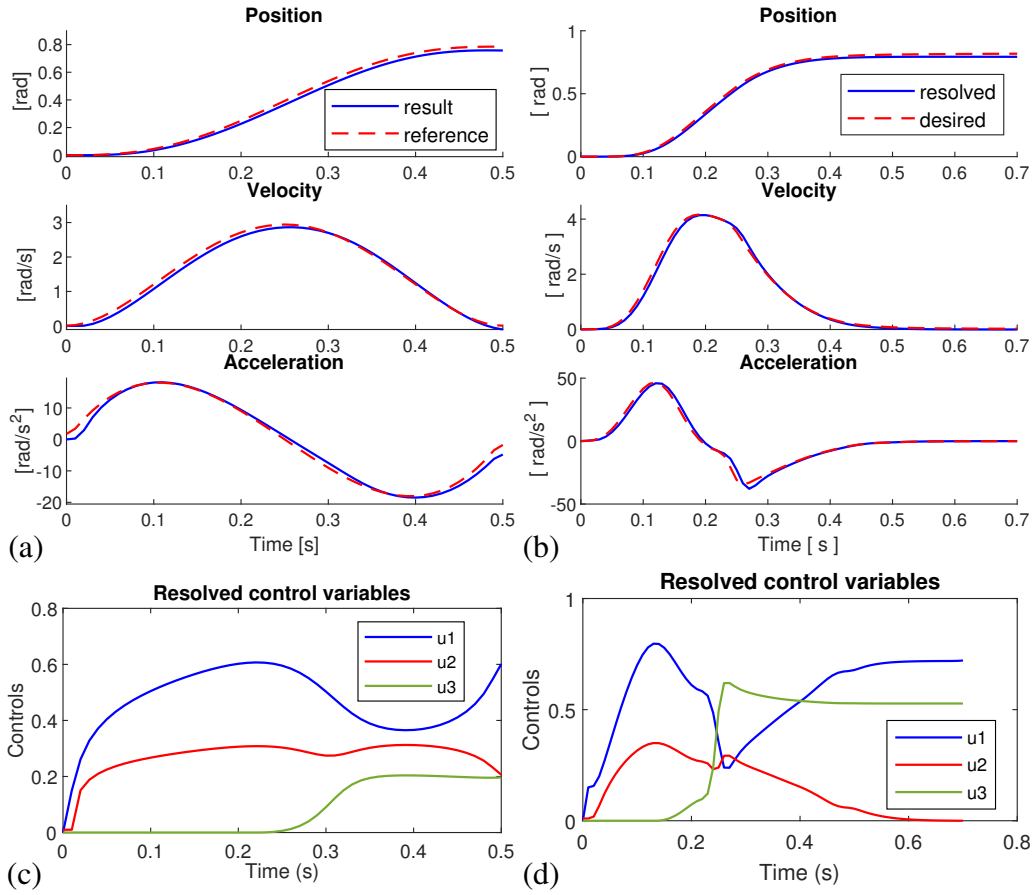


Figure 3.2: Example of joint trajectory tracking with torque feedback. (**Left**) Tracking a minimal jerk trajectory with gain $K_p = 10$. (**Right**) Tracking a trajectory given by optimal control with gain $K_p = 50$. The control u_1, u_2 are resolved to position command, u_3 is the damping command.

possibility of representing and memorizing the trajectory in the joint space or other compact forms, and dramatically reduces the computational burden for real-time control.

Another possible application of the null-space controller is in Imitation Learning. As pointed out in Chapter 2, variable physical impedance could be independently optimized to improve the task performance. The concept has been demonstrated in our previous work (Zhao et al. (2018)), where the damping profile was optimized in the null-space for an active impedance controlled robot Sawyer, thus contributing to further performance improvement. Concerning the energy effi-

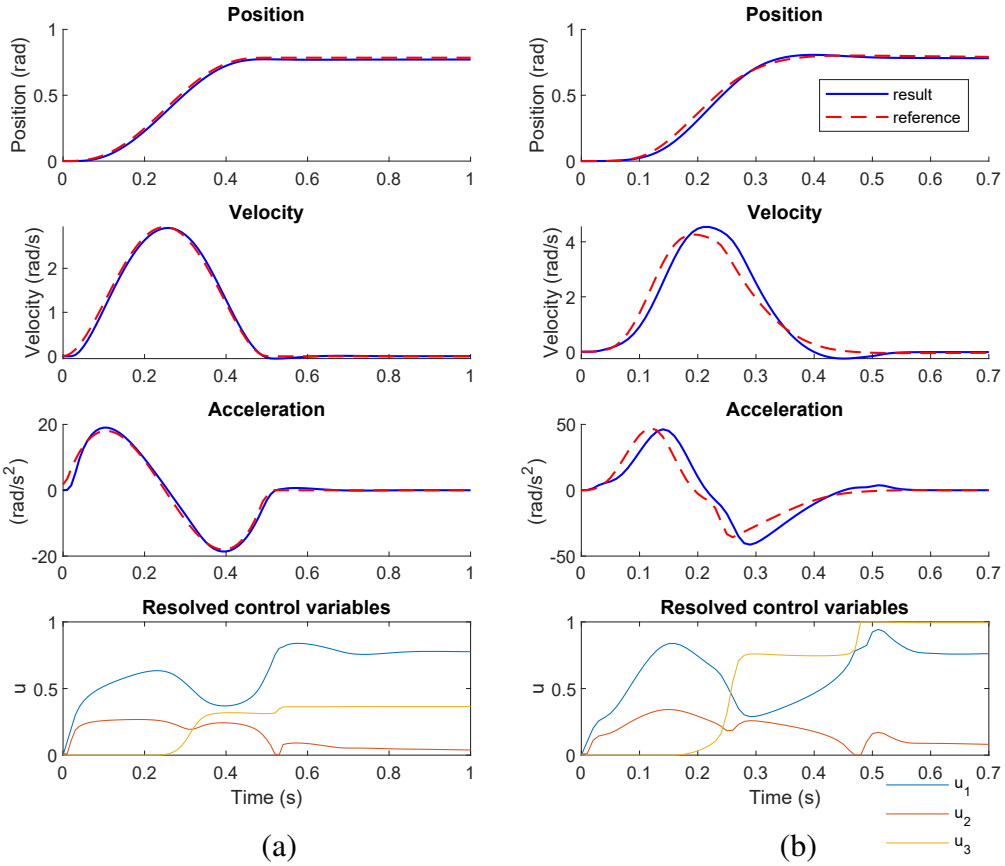


Figure 3.3: Example of joint trajectory tracking with joint feedback. (**Left**) Tracking a minimal jerk trajectory with gain $K_1, K_2, K_3 = 1000, 500, 50$. (**Right**) Tracking a trajectory given by optimal control with the same gain.

ciency problem, damping should be reduced when it is unnecessary to minimize energy dissipation. This intuition may lead to a bang-bang control strategy similarly as the maximum velocity problem shown by Haddadin et al. (2011b). In such a case, the optimal control needs to find a best switching time of the bang-bang strategy. When the actuator is required to find an optimal impedance control for tracking joint trajectory and or stiffness, the null-space controller may be optimized by formulating the closed form controller into the system dynamics as shown in Howard et al. (2013) and Zhao et al. (2018). By doing so, the optimal control problem becomes one of finding the optimum in the null space.

3.6 Summary

In this chapter we presented the ingredients of an optimal control framework. The modelling of system dynamics was introduced based on the Euler-Lagrange method, followed by an example of the MACCEPA actuator. A numerical algorithm ILQR was introduced that is capable of solving a broad class of optimal control problems computationally efficiently.

We also provided a closed form feedback controller to track trajectory in the joint space. It allows null-space control term to be exploited for secondary control objective. Since the formulation is quite general, it can be implemented for any actuators whose torque is non-linear in control.

The optimal control framework will be used throughout this thesis. While in Chapter 6 we will show how the closed form tracking controller can be used at low-level and enables high-level planning and learning.

Chapter 4

Energy Regeneration in Variable Impedance Actuators

In this chapter, we formally address the research question identified in Chapter 2, by proposing a regenerative damping control scheme specifically designed for the need of VIAs. By incorporating energy regeneration, the efficiency trade-off problem becomes more complex as the damping needs to balance its own functionality to harvest energy. We show that, via an optimal control approach, the trade-off can be balanced by taking into account of the regeneration estimation.

4.1 Introduction

In Chapter 2 we discussed the possibility of harvesting energy dissipated by the damper and recharging it to the power source. This chapter presents a circuit design that enables adjustment of the damping effect, while charging the current generated by bi-directional joint movements to a unidirectional power source (Wu and Howard (2018); Wu and Howard (2020a)).

The relation of the damping effect with the power of regeneration of the proposed damping module is investigated by analysis and physical experiment. The result shows that a non-monotonic relation emerges that requires balancing the trade-off between damping allocation and energy regeneration in a non-trivial way. Nevertheless, the proposed controllable damping module is evaluated in

terms of movement performance and energy recovery of a simple reaching task on an ideal pendulum model and a more realistic VIA model, and shows that an optimal trade-off is achievable.

4.1.1 Dynamic and regenerative braking

Among the different methods of implementing variable physical damping in VIAs, *damping by motor braking* presents the greatest promise for incorporating energy harvesting by utilizing the regenerative braking technique. For this, two main approaches are available, namely (i) *dynamic braking* and (ii) *regenerative braking*. In both cases, the back electromotive force is used to resist movement proportional to the effective resistance of the damper motor circuit, causing a variable damping effect. This damping effect is due to the electrical modulation and thus does not take account of the mechanical damping, which is mainly the friction in the gearbox of the damping motor.

Scheme 1 - Dynamic braking

Dynamic braking in the context of VIA design was first proposed by Radulescu et al. (2012). A circuit diagram for this scheme is illustrated in Figure 4.1(a). In this mode, the damping effect is modulated by changing the duty-cycle D_d that controls the portion of time that a switch S_1 spends in the open or closed position, thereby altering the effective resistance of the circuit. The damping coefficient follows the equation

$$d = \frac{n_d^2 k_t^2 D_d}{R_m} = \bar{d_1} D_d \tag{4.1}$$

where n_d is the gear ratio of damping motor, k_t is the torque constant and R_m is the resistance of the motor. Note that, since $0 \leq D_d \leq 1$, the maximum damping coefficient that can be provided by dynamic braking is $\bar{d_1} = n_d^2 k_t^2 / R_m$.

In energy terms, dynamic braking is effective since it dissipates kinetic energy of the output link as heat in the electrical circuit. It does not, however, charge energy to any electrical source, so the regeneration power is zero ($P_{\text{rege}} = 0$). In other words, this (potentially useful) energy is simply discarded, reducing the overall energy efficiency of the system.

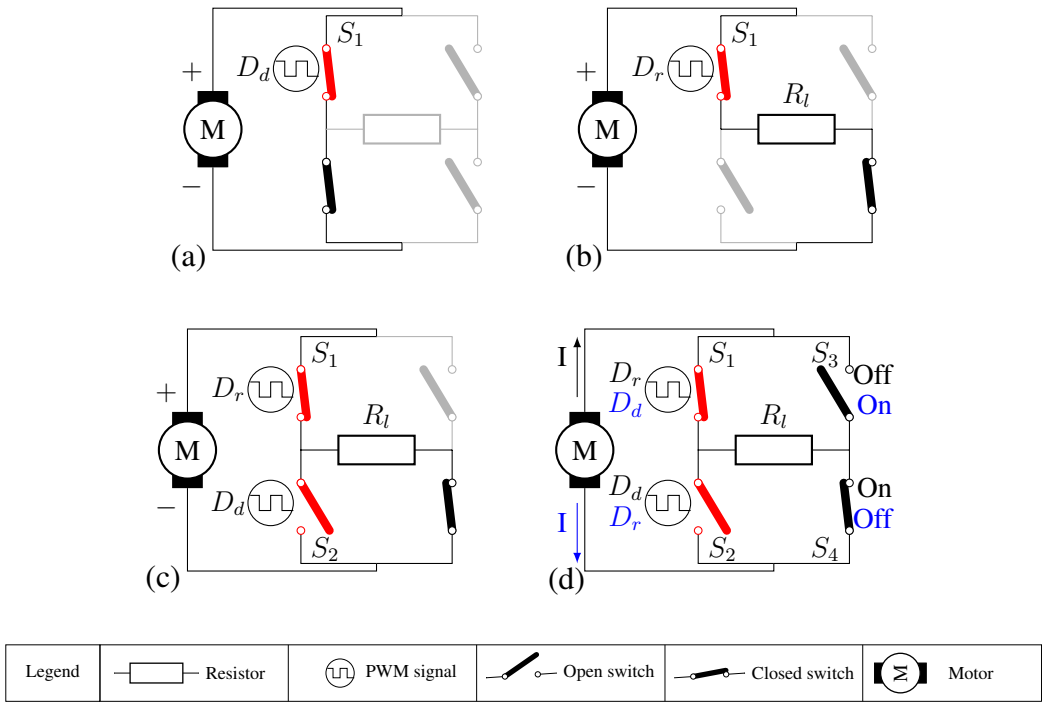


Figure 4.1: Conceptual diagrams of (a) dynamic, (b) regenerative, (c) hybrid dynamic-regenerative, and (d) bidirectional hybrid dynamic-regenerative braking circuits. In the circuits, $S_i, i = 1, ..., 4$ are switches and D_r, D_d are pulse-width modulated (PWM) control signals. The switches are coloured in red when they are actively controlled by PWM signals and in black when they are permanently open or closed. The diagrams are sketched in a synthesized way to show how the different functions can actually be realized in a single circuit (in diagrams (a)-(c), the parts of the circuit shown in grey should be considered disabled or disconnected). Generally, the circuit consists of an energy storage element (*e.g.,* a battery or supercapacitor), whose dissipation is modelled by a resistor R_l, connected to a DC motor as a damper, with a switching mechanism. In diagram (d), the current I can have two directions indicated by black and blue arrows. When it flows in the direction indicated by the black arrow, S_3 is set to off and S_4 is on; the PWM signals for S_1, S_2 are control variables D_r, D_d respectively. When the current flow is reversed in the direction of blue arrow, the control variable used for S_1 is switched to D_d and D_r controls S_2; the on-off modes of S_3, S_4 are switched as well.

Scheme 2 - Regenerative braking

Regenerative braking refers to the situation where the power generated by the motor through kinetic motion of the output link is used to recharge an electrical

storage element (*e.g.,* battery, supercapacitor). To implement regenerative braking, the electrical storage element can be simply connected to the circuit of the damping motor, as shown in Donelan et al. (2008). In the context of VIA design, this can be implemented through the circuit in Figure 4.1(b).

In regenerative braking mode, the damping effect is dependent on the combined effective resistance of the circuit containing the electrical storage element. Similar to dynamic braking, this can be modulated by controlling the duty-cycle D_r of a switch. The damping coefficient and the regeneration power can be calculated as

$$d = \frac{n_d^2 k_t^2 D_r}{R_m + R_l} = \bar{d}_2 D_r \tag{4.2}$$

and

$$P_{\text{rege}} = \frac{R_l n_d^2 k_b^2 \dot{q}^2 D_r}{(R_m + R_l)^2} = \alpha \bar{d}_2 \dot{q}^2 D_r, \tag{4.3}$$

respectively, where R_l is the internal resistance of the electrical storage element (*e.g.,* a battery), $\alpha = R_l/(R_m+R_l)$ and $\bar{d}_2 = n_d^2 k_t^2/(R_m+R_l)$. k_b is the back-EMF constant and is equal to k_t.

Note that, introducing regenerative braking means that the mechanical energy that is otherwise discarded in the dynamic braking scheme can be harvested, enhancing the overall energy efficiency of the system. However, note also that, compared to dynamic braking, the maximum damping coefficient $\bar{d}_2$ that can be produced by regenerative braking, is *decreased* since adding an electrical load for charging increases the total equivalent resistance of the circuit. This can be a drawback in applications where higher levels of damping are needed (*e.g.,* when there is need for a high dynamic response and therefore heavy braking of rapid movements). Another issue is that, since the electrical storage element is usually unidirectional, the current in Scheme 2 has to be uni-directional as well. In order to deal with current following in both directions resulted from the bidirectional movements (which is common in robotic applications), a reversing mechanism is needed. One solution for this is to introduce another switch, as illustrated in Figure 4.1(d) — this will be introduced in detail in §4.2.

4.2 Hybrid dynamic-regenerative braking

To meet the requirements for the regenerative damping system to be used with VIAs, this section presents a variable damping scheme — termed *hybrid braking* — that switches between dynamic braking (Scheme 1 as described in §4.1.1) and pure regenerative braking (Scheme 2 in §4.1.1) to achieve the benefits of both. Note that, the proposed scheme is designed as a standalone module to be used in conjunction with a variable stiffness mechanism, which is not restricted to the example taken in this thesis (where the joint is connected in series with a variable spring to a motor, and the variable spring is controlled by a second motor to adjust the joint stiffness), but could be agonistic/antagonistic employment of springs, series elastic actuators (SEAs), or other complex arrangement that cannot be represented by the model in Figure 2.3. It could, in theory, also be applied to any joint with high compliance or backdrivability, even when the drive motor is rigidly connected to the joint.

4.2.1 Hybrid damping circuit

The hybrid damping scheme proposed here is implemented through the circuit depicted in Figure 4.1 (c). It uses two switches (denoted S_i, $i \in \{1, 2\}$) that switch at high frequency between (i) pure regenerative braking, and (ii) a blend of dynamic and regenerative braking. The principle by which the proposed scheme operates is as follows.

When switch S_2 is open, the module acts in regenerative braking mode, whereby current flows through the power storage element, with the effective resistance (damping level) determined by the duty cycle of S_1. (Note that, this results in an equivalent circuit to that used in Scheme 2, *cf.* Figure 4.1 (b).) On the other hand, when S_1 and S_2 are closed, there is a short circuit that causes current to bypass the resistive load R_l, creating a dynamic braking effect. In this case, the damping level can be determined by keeping S_1 closed and modulating the duty cycle of S_2. This enables a third braking scheme to be defined, alongside Schemes 1 and 2, as follows.

Scheme 3 - Hybrid braking

When the required damping d^* is small enough, *i.e.,* $d^* \leq \bar{d}_2$, it can be provided by pure regenerative braking, so S_2 is opened ($D_d = 0$). When the required damping is greater, *i.e.,* $d^* > \bar{d}_2$, S_1 is closed ($D_r = 1$) and D_d is used to control S_2 to blend dynamic and regenerative braking.

The resulting damping coefficient and regeneration power are:

$$d = \bar{d}_2 D_r + \alpha \bar{d}_3 D_d \tag{4.4}$$

$$P_{\text{rege}} = \alpha \bar{d}_2 \dot{q}^2 (D_r - D_d) = P^0_{\text{rege}} \dot{q}^2 \tag{4.5}$$

where $P^0_{\text{rege}} = \alpha \bar{d}_2 (D_r - D_d)$ is the regeneration coefficient. Note that, if $D_r = D_d = 1$, the same maximum damping coefficient as that achievable in a pure dynamic braking can be achieved, *i.e.,* $\bar{d}_3 = \bar{d}_1$. This, however, comes at the cost of the regeneration power vanishing ($P_{\text{rege}} = 0$).

4.2.2 Hybrid Damping Control Modes

In principle, each of the switches in the proposed circuit may be independently controlled by its own duty-cycle. While this enhances the flexibility of the damping module, it also introduces an undesirable layer of complexity to its control.

To address this, and enable the simple control of the module through a single control variable $u \in [0, 1]$, the duty cycles of the switches can be coupled through the following relation

$$D_r = \begin{cases} \dfrac{u}{u_r}, u \leqslant u_r \\ 1, u > u_r \end{cases}$$

$$D_d = \begin{cases} 0, u \leqslant u_r \\ \dfrac{u - u_r}{1 - u_r}, u > u_r \end{cases} \tag{4.6}$$

where u_r corresponds to the maximum damping coefficient of regenerative braking ($d(u_r) = \bar{d}_2$) and depends on the user's selection. In this work, u_r is chosen to be 0.5. Substituting (4.6) into (4.4), the damping coefficient as a function of u

is simplified to

$$d(u) = \bar{d}_3 u. \tag{4.7}$$

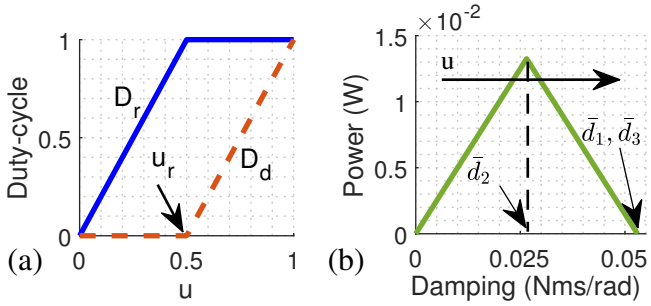


Figure 4.2: Hybrid damping control modes. (a) Mapping from control input u to duty cycles D_r, D_d. (b) Relation between regeneration power (P_{rege}) and damping.

As illustrated in Figure 4.2(a), when $u \leqslant 0.5$, D_d remains at zero (*i.e.,* switch S_2 is open) and D_r is linearly mapped from $u \in [0, 0.5]$ to $[0, 1]$, while when $u > 0.5$, D_r is held at unity ($D_r = 1$ so S_1 is closed) and D_d is linearly mapped from $u \in [0.5, 1]$ to $[0, 1]$.

The relation between the damping coefficient d and the power regeneration P_{rege} for a fixed angular velocity is shown in Figure 4.2(b). As can be seen, the relationship is non-monotonic and there is a peak for P_{rege} when $d = \bar{d}_2$, *i.e.,* at the upper boundary of the pure regenerative braking domain.

4.2.3 Bidirectional damping

The hybrid damping circuit described so far enables the modulation of damping force associated with unidirectional motion of the output link. In order to realize damping of bidirectional motion (as is common in many robotic applications), it is necessary to ensure that the current generated by the damping motor always flows into the positive terminal of the electrical storage element. This can be achieved by a four-switch design of the damping circuit, as illustrated in Figure 4.1(d). When the current flows from the positive terminal of the damping motor (as shown by the black arrow in Figure 4.1(d)), S_3 is open and S_4 is switched on. When the current flows from the negative terminal of the motor (as shown by the grey arrow), S_3 is

closed and S_4 is open, and S_1 is controlled by D_d and S_2 is controlled by D_r.

It should be further noted that, this latter circuit, implements the (bidirectional versions of) the two damping schemes outlined in §4.1.1 as special cases. For example, (i) holding S_2, S_3 open, S_4 closed and varying the duty cycle of S_1 results in regenerative braking, while (ii) holding S_3, S_4 open, S_1 closed, and varying the duty cycle of S_2 results in pure dynamic braking. In other words, the same hardware can be used to realize all three damping schemes. In the following sections, for brevity, the term regenerative damping will be used to refer to the proposed hybrid damping scheme in the context of VIAs.

4.2.4 Design choices

From (4.3) and (4.5), it can be seen that the maximum amount of regeneration power is proportional to the squared gearbox ratio n_d^2 and squared torque constant k_t^2. This suggests that one should select electric motors that have higher torque constant and gearboxes with higher gear ratio to increase the maximum regeneration power. However, higher gear ratio tends to increase the friction and reflected inertia. In general, a DC motor and gearbox that has higher torque constant, lower friction and lower inertia is preferred. The gear ratio can be decided by considering the desired range of damping effect and by trial and error.

4.2.5 Physical realisation of the damping module

This section presents the physical realization of the hybrid damping circuit design introduced above and an experiment to verify the theoretical predictions about the damping/regeneration performance trade-off. The experimental set up is shown in Figure 4.3. As a simple test-rig, two identical DC motors (Maxon A-max 22/110125) are coupled through a pair of spur gears to enable one motor (driver) to drive the other (damper), see Figure 4.3(b) and (c). The two motors have the same gearhead with $n_d = 20$. The torque constant is $k_t = 0.0212\mathrm{Nm/A}$ and the motor resistance $R_m = 21.2\Omega$.

The damper motor is connected to the circuit depicted in Figure 4.3(a) (via the nodes P1 and P2), that is the physical realization of the conceptual diagram Figure 4.1(d). In this circuit design, a pair of N-channel MOSFETs (IRF520) is

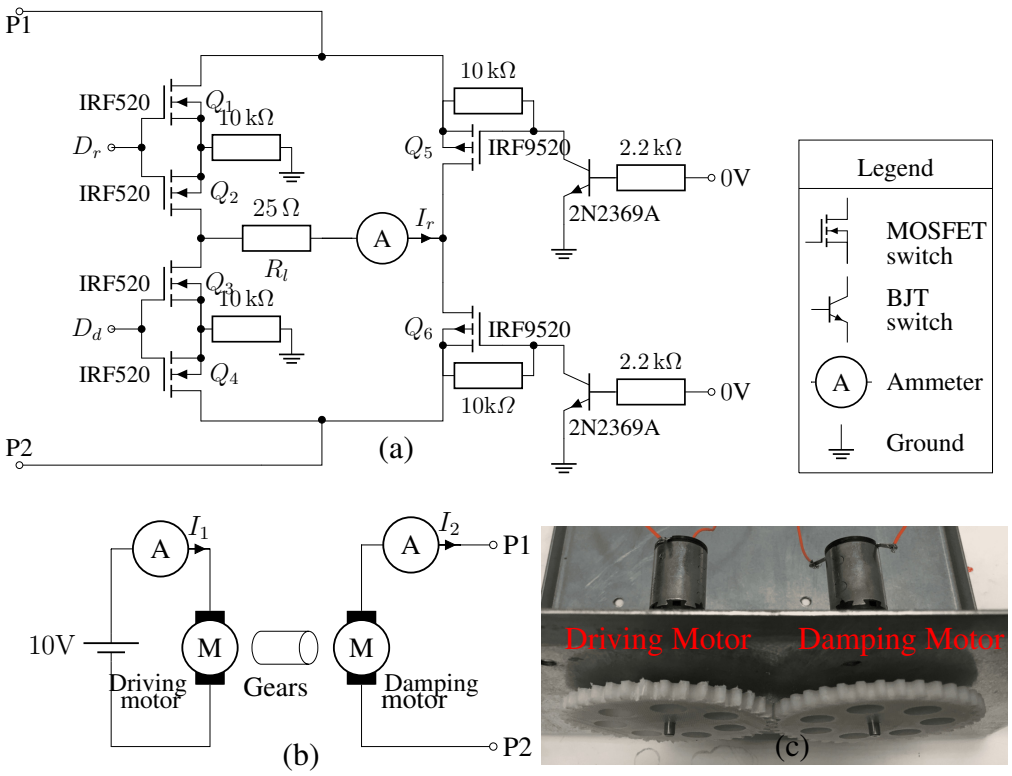


Figure 4.3: Damping and power regeneration measurement experiment setup. Shown are (a) circuit diagram of the damping module, (b) diagram of the experiment setup, and (c) the test rig. The terminals of the damping motor are connected to nodes P1 and P2 of the control circuit.

used as one switch to make sure that the switching mechanism works properly for bidirectional current. In Figure 4.3(a), the pair of Q_1, Q_2 works as the switch S_1, and Q_3, Q_4 make the switch S_2. Two P-channel MOSFETs (IRF9520) with BJTs (2N2369A) are used as switches S_3, S_4. The duty-cycles D_r, D_d are controlled by PWM signals from an Arduino Mega2560 board. By setting 0V signals on the control pins for Q_5, Q_6, they are open for just one current direction but closed for the other. For the ease of power measurement, a resistor is used to represent the electrical load ($R_l = 25.3\Omega$).

In the experiment, the driving motor is used to drive the system while the damping applied by the second motor is varied, and the resultant motion (motor speeds and energy regeneration) is recorded. Specifically, the driving motor is powered at 10V ($V_{bb} = 10\text{V}$) constantly by a laboratory DC power supply while

the damping motor control input u is varied from 0 to 1 in increments of 0.1 (with the corresponding duty-cycles D_r, D_d computed by (4.6)). Simultaneously, three multimeters (Rapid DMM 318) are used to measure the currents I_1, I_2, I_r through the driving motor, damping motor and the electrical load R_l (represented by a resistor) respectively. The latter data are used to compute the angular speed of the motors ω and the damping torque τ_d according to

$$V_{bb} = I_1 R_m + n_d k_t \omega \tag{4.8}$$

$$\tau_d = n_d k_t I_2 = d(u)\omega. \tag{4.9}$$

The damping coefficient $d(u)$ for a given u is then estimated by

$$\hat{d}(u) = \frac{n_d^2 k_t^2 I_2}{V_{bb} - I_1 R_m} \tag{4.10}$$

and the regeneration coefficient P^0_{rege} is estimated as

$$\hat{P}^0_{\text{rege}} = \frac{n_d^2 k_t^2 I_r^2 R_l}{(V_{bb} - I_1 R_m)^2}. \tag{4.11}$$

The results based on the data collected from 10 repetitions of the experiment is plotted in Figure 4.4 alongside the theoretical predictions (from §4.2.2).

It can be seen that, the experimental data (blue lines in Figure 4.4) is in good agreement with the model predictions (red lines in Figure 4.4). By increasing u from 0 to 1, the damping coefficient d increases almost proportionally. Furthermore, it is verified that, when fixing the angular speed (P_{rege} has been normalized to estimate P^0_{rege}), the relation between P^0_{rege} and u is non-monotonic with a peak found at $u = 0.5$. However, the experimental data indicates that for both regions ($u \in [0, 0.5]$ and $u \in [0.5, 1]$), the regeneration coefficient is not linearly dependent on the control input. This modelling error might be due to unmodelled effects such as circuit inductance and switching frequency.

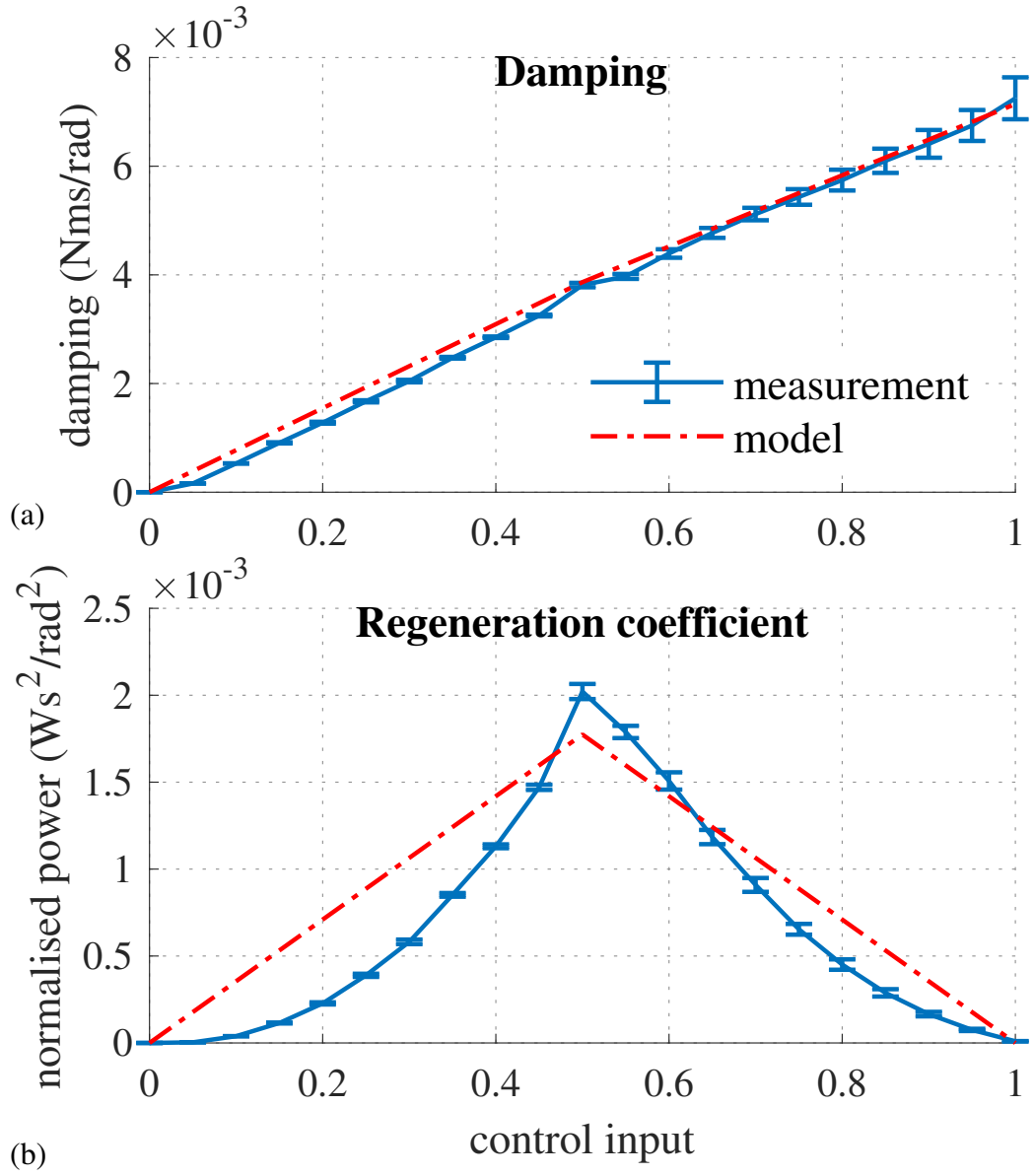


Figure 4.4: Results of the damping test experiment. The (a) damping coefficient, and (b) regeneration coefficient for each tested control input $u \in [0, 1]$ are shown. The blue error bars represent the means and standard deviations of data points for 10 repetitions of the experiment and the red line shows the values predicted by the model.

4.3 Simulation and evaluation with ideal VIA

Having verified the feasibility of physically implementing the proposed damping module, it is necessary to evaluate its use in the context of robot control. As

noted in §2.4.1, the dual role of the damping module, both for braking and energy harvesting introduces a trade-off between task performance and energy efficiency. To resolve this, it is proposed to employ optimal control to determine the best damping modulation strategy according to task demands.

This section evaluates the proposed hybrid braking scheme in comparison to pure dynamic or regenerative braking through numerical simulation of (i) a simple pendulum actuated with an ideal VIA, and (ii) a more realistic simulation of a physical VIA, namely the MACCEPA-VD introduced in §3.3.2.

4.3.1 Simple pendulum with ideal VIA

The aim of the first evaluation is to illustrate the effectiveness of the hybrid braking scheme in the context of a simple example task of target reaching.

For this, a model of a simple pendulum, subject to viscous friction and actuated by an ideal VIA is used

$$ml^2\ddot{q} + b\dot{q} = k(u_2)(u_1 - q) - d(u_3)\dot{q}. \tag{4.12}$$

Here, for simplicity, $m = 1\text{kg}, l = 1\text{m}, b = 0.01\text{Nms/rad}$. The motor positions θ_1, θ_2 are assumed to be directly controlled by control inputs u_1, u_2. u_3 is the control input for damping d. The right-hand side of (4.12) is the joint torque applied by the ideal VIA, $u_1 \in [-\pi/2, \pi/2]\,\text{rad}$ controls the equilibrium position and the stiffness $k(u_2)$ is proportional to the control input $u_2 \in [0, 1]$, ***i.e.,***

$$k(u_2) = \bar{k}u_2, \tag{4.13}$$

where $\bar{k} = 200\,\text{Nm/rad}$ is the maximum stiffness. The damping $d(u_3)$ as a function of u_3 is given by (4.7). The corresponding power of regeneration P_{rege} is assumed to be computed by the model introduced in §4.2. The parameters[1] that characterize the variable damping module are selected to be $\bar{d}_3 = 50\,\text{Nms/rad}$, $\bar{d}_2 = 25\text{Nms/rad}$ and $\alpha = 0.5$. The control frequency is set to 50Hz.

The task is to reach a target $q^* = \pi/3\,\text{rad}$ from the initial position $q = 0\,\text{rad}$

[1]These parameters are arbitrarily chosen to give response within a second. Experimentation shows the result is not sensitive to these values.

within a finite time t_f as quickly and accurately as possible, while minimizing the energy consumption and control effort. This can be described through minimization of the cost function[2]

$$J = \int_0^{t_f} [w_1(q(t) - q^*)^2 + w_2(u_1(t) - q^*)^2 + w_3 u_2^2(t) - w_4 P_{\text{rege}}] \, \mathrm{d}t, \quad (4.14)$$

where $w_1 = 1000$, $w_2 = w_3 = 1$, $w_4 = 0.01$ are weighting parameters. These parameters are selected to take account of the different scales of the terms and allow reaching within a second. In the cost function, the first term represents the reaching accuracy and drives the plant to reach the target quickly; the second term is used to penalize deviation of the equilibrium position from the target to increase stability; the third term encourages using lower pretension (corresponding to lower stiffness); and the fourth term is used to encourage using the regeneration.

To simplify the analysis, in the below, the command for equilibrium position is fixed at $u_1 = \pi/3$, while the commands for stiffness and damping are allowed to vary. The optimal open-loop control sequence for the latter is computed through the ILQR (ILQR) method (Li and Todorov (2004)) with the proposed hybrid braking scheme, and the resultant trajectory of the system is computed by simulating the execution of the open-loop command using the 4th Order Runge-Kutta method.

To evaluate the energy efficiency of the proposed approach, the total mechanical work[3] and the total regenerated energy are computed from the resultant trajectories, *i.e.,*

$$E = \int_0^{t_f} k(u_1 - q)\dot{q} \, \mathrm{d}t \quad (4.15)$$

$$E_{\text{rege}} = \int_0^{t_f} P_{\text{rege}}(t) \, \mathrm{d}t, \quad (4.16)$$

[2]The use of regeneration power as in the cost function has been justified by our preliminary study, see Appendix C.

[3]As motors are not explicitly involved in the model, the mechanical work computed here corresponds with the integration of the power delivered onto the plant (P_{out} in Figure 2.5), not the power from motors side (P_{in}).

respectively. The net energy cost can be defined as

$$E_{\text{net}} = E - E_{\text{rege}}. \tag{4.17}$$

The percentage ratio of energy regeneration [4] can be computed by

$$\eta = \frac{E_{\text{rege}}}{E}. \tag{4.18}$$

For comparison, the experiment is repeated with (i) pure dynamic braking (Scheme 1), (ii) pure regenerative braking (Scheme 2), (iii) the case where the damping is fixed at the maximum power of regeneration ($d = \bar{d}_2$), and (iv) a critically damped system. In the latter, the stiffness is chosen to be $k = 100\text{Nm/rad}$ and the damping is fixed to $d = 20\text{Nms/rad}$ such that the damping ratio $\zeta = d/2\sqrt{km} = 1$.

The results are illustrated in Figure 4.5. As can be seen, the trajectory of the critically damped system reaches the target slowly but without overshoot (Figure 4.5(a)). The system with fixed damping reaches the target quicker than the critically damped one, because it can exploit the variable stiffness. The system with regenerative braking reaches the target quicker still, however, since the damping range is limited in this case, it suffers from overshoot once it reaches the target. In contrast, the dynamic braking and hybrid braking systems reach the target quickest without overshot, so perform best in terms of accuracy.

Looking at Figure 4.5(c), however, it can be observed that the dynamic braking performs worst in terms of net energy cost, since no energy is recovered throughout the movement. This contrasts with the hybrid approach, that achieves fast and accurate movement while also achieving 27.4% energy recovery, thereby lowering the net energy cost.

Overall, the proposed hybrid scheme offers a good trade-off between task accuracy and energy efficiency.

[4] Note that, for simplicity, it is assumed here that there is 100% kinetic to electric energy transmission efficiency of the DC motor. In practice, losses are likely to occur due to friction and losses in the conversion from the mechanical to the electrical domain.

4.3.2 Optimal reaching with the MACCEPA-VD

To evaluate the proposed scheme on a more realistic variable impedance actuation system, the MACCEPA-VD mechanism (Figure 3.1) on page 50, as introduced in §3.3.2, is chosen as an example. In the MACCEPA-VD, the equilibrium position

Symbol	Value	Unit	Meaning
B	3.6	cm	length of lever arm
C	13.5	cm	pin displacement
r	1.5	cm	radius
κ	394	N/m	spring constant
m	0.0015	kgm^2	inertia
b	0.0023	$\mathrm{Nms/rad}$	friction coefficient

Table 4.1: MACCEPA-VD model parameters.

and joint stiffness are controlled by a servomotor each respectively, while the damping coefficient is modulated by changing the duty cycle of the circuit of a third DC motor attached rigidly to the joint. The model parameters are specified in Table 4.1.

In addition, the motor angles θ_1, θ_2 are controlled by $u_1 \in [-\pi/3, \pi/3], u_2 \in [0, \pi/3]$ respectively. The servomotor dynamics (3.29), (3.30) are assumed to behave as a critically damped system, with $\beta = 30$. The damping coefficient $d(u_3)$ depends on control input u_3 and is calculated according to the damping scheme used (*i.e.,* (4.1), (4.2) or (4.7)).

To evaluate the proposed hybrid damping method, ILQR is used to determine the optimal open-loop control sequence for the task of reaching a target point $q = \pi/4\,\mathrm{rad}$ starting from initial state [5] $\mathbf{x}_0 = (0, 0, 0, 0, 0, 0)^\mathsf{T}$ within finite time $t_f = 2\,\mathrm{s}$, using the proposed hybrid damping scheme (see §4.2). The cost function for optimization takes the same form as (4.14), where weighting parameters are $w_1 = 10^3, w_2 = w_3 = 10^{-4}, w_4 = 10^{-6}$.

For comparison, the experiment is repeated using dynamic and regenerative braking (Schemes 1 and 2), a fixed damping coefficient of $d = \bar{d}_1$, and a 'critically damped' system in which, following Radulescu et al. (2012), the instantaneous

[5] Recall Chapter 3, the state vector is defined as $\mathbf{x} = (q, \dot{q}, \theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2)^\mathsf{T}$.

damping ratio is held at $\zeta(t) = 1$ by enforcing the stiffness-damping relationship $d(t) = 2\sqrt{k(t)m}$.

The results are shown in Figure 4.6. There, it can be seen that, the 'critically damped' system avoids overshoot, but reaches the target slowly. It also has the lowest energy consumption, in part due to its sluggish response in moving to the target. The fixed damping trajectory reaches the target faster, but slower than the trajectories using dynamic and hybrid schemes. Although it regenerates the most energy, it also incurs the highest cost in terms of mechanical energy, so overall the net energy cost is higher than Schemes 2 and 3 (see Figure 4.6(c)).

The regenerative braking scheme has better energy efficiency as its net energy cost is lower but it suffers greatest overshoot due to its restricted damping range, while the dynamic braking scheme achieves higher accuracy but at the expense of higher net energy cost. The former has higher total mechanical work, but regenerates 41.0% and results in lower net energy cost. In comparison, hybrid braking achieves almost identical performance in terms of accuracy, but at higher energy efficiency.

The damping commands of both examples, as illustrated in Figure 4.5(d) and Figure 4.6(d), show unnecessary oscillations after the joint reaches the targets. This is because the fourth term in the cost function (4.14) P_{rege} remains zero when the joint velocity $\dot{q} = 0$, regardless of the value of the damping. In order to eliminate the oscillations, an extra quadratic regularization term u_3^2 can be added to the cost function. Alternatively, one can use $(u_3 - 0.5)^2$ to encourage keeping the damping around the level that produces maximum regeneration coefficient and regularize the oscillation in the meanwhile.

An oscillation of the stiffness profile can be observed in the MACCEPA-VD example (Figure 4.6(b)). An interpretation from the perspective of optimal control is that when the joint reaches the target it does not come to a perfect equilibrium state ($q = q^*, \dot{q} = \ddot{q} = 0$) immediately. Thus the EP and stiffness motors are still manipulated after reaching the target and exhibits the "lead and lag" behaviour to achieve a stable state of the system. This explains that why the oscillation of stiffness is absent in the toy example (Figure 4.5(b)). Since the EP motor command is fixed at the target position, the stiffness motor cannot be modulated alone to apply positive and negative torque on the joint repeatedly after reaching the

target. In this case, the optimal strategy has to rely on controlling the damping to brake the movement. A possible way to decrease the oscillation of stiffness in the MACCEPA-VD example is simply increase the weights w_2, w_3 of the quadratic terms of u_1 and u_2 in the cost function. Also, one can add more regularization terms such as $\dot{u}_1^2, \dot{u}_2^2$ to penalize the oscillation.

4.4 Conclusion

This chapter proposed an extension to variable damping module design for VIAs based on the motor braking effect. In contrast to previous, pure dynamic braking designs, the proposed approach provides a solution for realizing controllable damping, which enables the VIAs to regenerate dissipated energy from bidirectional movement to charge a unidirectional electric storage element. Furthermore, it overcomes the drawback of a reduction in the maximum damping effect found in pure regenerative braking schemes.

The control input for this damping module simply varies from 0 to 1, representing a proportional percentage of the maximum damping. As the power regeneration has a non-monotonic relation with the control input and damping coefficient (as verified by experiment), the balancing between damping allocation and energy regeneration needs to be treated with care. However, application of the hybrid damping module to VIAs verified by experiments, shows that the actuation redundancy is solved by optimal control successfully to achieve fast smooth movement while still enabling power regeneration.

Furthermore, an experiment is presented in which the damping module is realized in hardware, verifying the theoretical predictions about the damping module's behaviour.

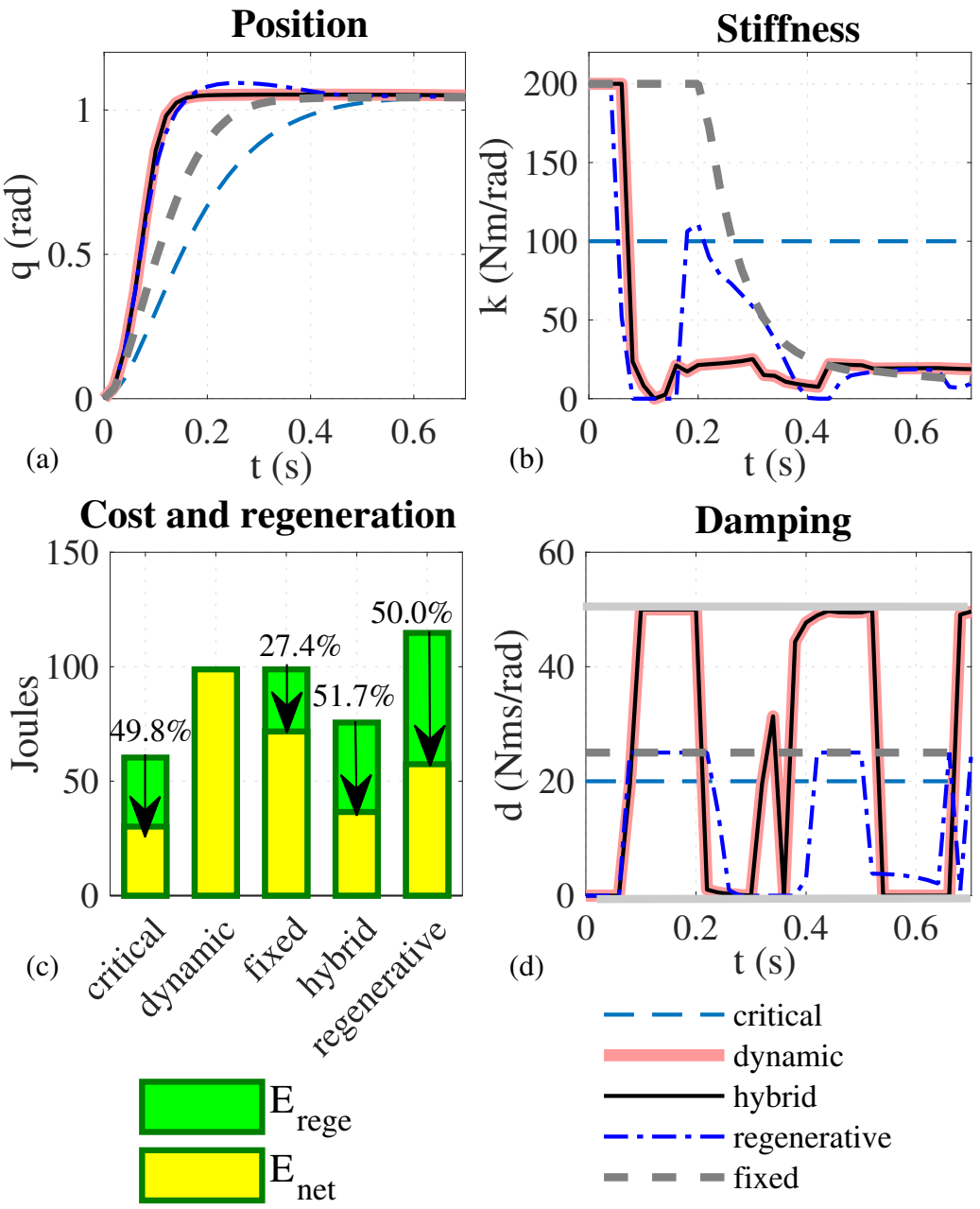


Figure 4.5: Test of reaching task on a simple pendulum with ideal VIA. Shown are optimal (a) joint angular trajectories, (b) stiffness profiles, (c) total mechanical work and percentage ratio of energy regeneration, and (d) damping profiles for different damping schemes. The proposed hybrid damping scheme is compared with (i) critical damping, (ii) pure dynamic braking, (iii) pure regenerative braking, and (iv) fixed damping ($d = \bar{d}_2$).

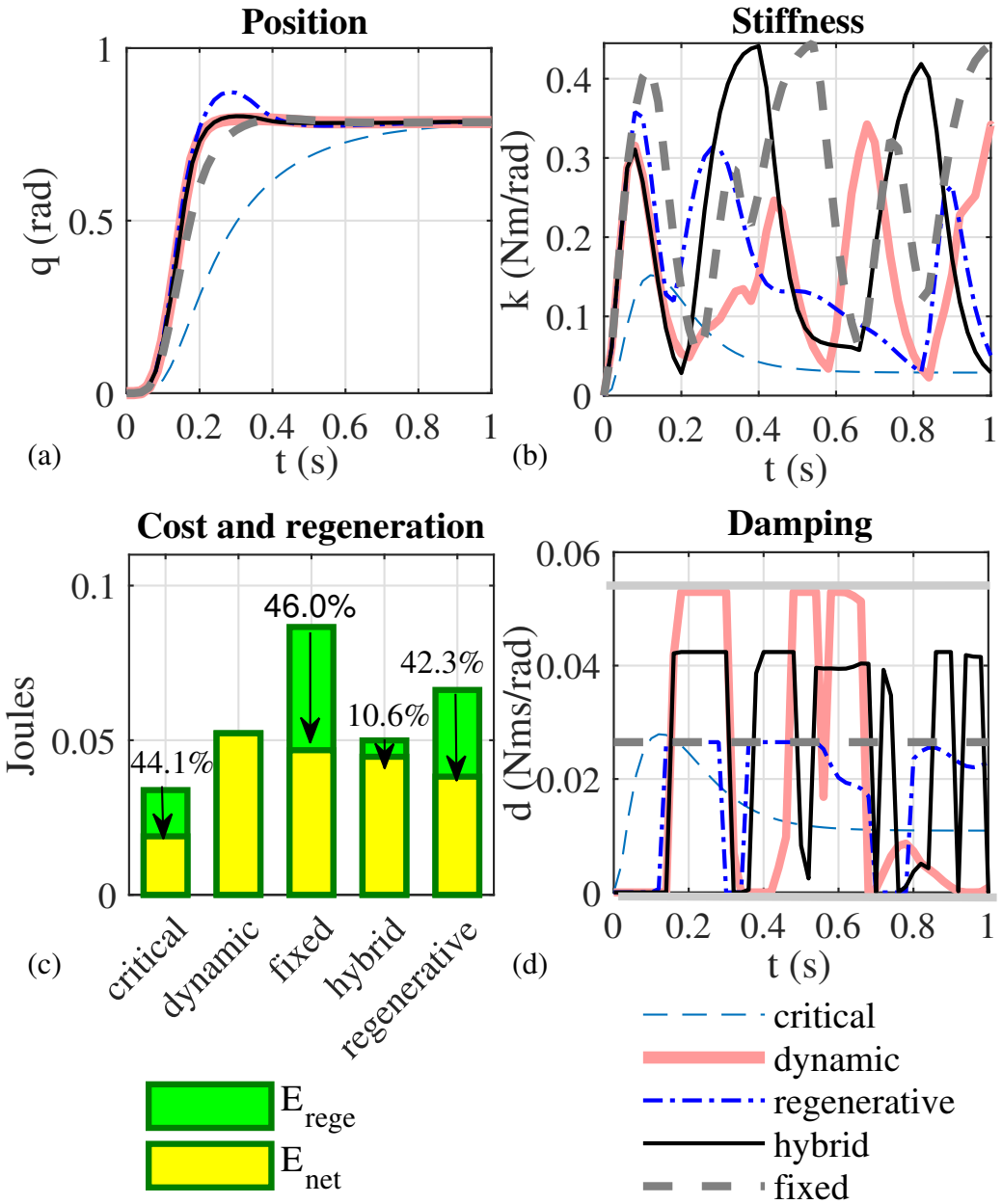


Figure 4.6: Test of reaching task with the MACCEPA-VD. Shown are optimal (a) joint angular trajectories, (b) stiffness, and (c) total mechanical work and percentage ratio of energy regeneration for different damping schemes, and (d) damping profiles.

Chapter 5

Energy Efficiency of Long-term Operation

In this chapter, with the regenerative damping module developed in previous chapter, we are ready to implement and test it on a real actuator. We will integrate it with a MACCEPA-VD and test its functionality. The questions we are interested in are how should we measure and evaluate the energy efficiency and how the introduced energy regeneration influences the efficiency trade-off? We will answer these questions by empirical study in the context of long-term operation.

5.1 Introduction

This chapter implements the proposed regenerative damping module on a physical robot driven by a VIA, to gain deeper insight into the role of variable regenerative damping and investigate the energy efficiency problem in the context of long-term deployment and operation of compliant robots. In contrast to the cyclic movement tasks commonly explored in prior work, this paper presents experiments in performing a stochastic movement task that mimics long term industrial operation and measures the performance of VIAs designed for versatile purposes. The results demonstrate that variable regenerative damping, *in combination with an optimally exploited variable stiffness mechanism*, can contribute to both enhanced dynamic performance and improved energy efficiency (in terms of both consump-

tion and regeneration). Measuring performance through four metrics (accuracy, settling time, energy consumption and regeneration), results reported here indicate that this approach can outperform schemes where stiffness and/or damping are fixed by up to 25%.

5.2 Long-term robotic deployment

In the real-world deployment of compliant robotic systems, VIAs must be able to withstand many use cycles with unpredictable task demands. This section reports an experiment designed to evaluate the likely effectiveness of the proposed damping scheme in the context of long-term use. For simplicity, the test case chosen is the task of performing consecutive point-to-point reaching movements to a series of random targets generated on the fly (*i.e.,* as the robot is moving). The aim is to examine the performance of the proposed scheme as compared to competing ones *where stiffness and/or damping are fixed* against multiple performance metrics. Note that, while such tasks are common in many robotic applications (*e.g.,* a robot deployed to tidy a room may have to reach and grasp many objects at uncertain locations), they are challenging from the point of view of energy management, since the movements are non-periodic and unpredictable . The following reports the experimental design and procedure in detail.

5.2.1 Hardware specifications

To evaluate its use, the damping module developed in §4.2.5 is implemented on a physical VIA. Specifically, the experimental platform consists of a 3D-printed, single-joint robot using the MACCEPA-VD mechanism for actuation, where the joint stiffness is adjusted by changing spring pretension (see Figure 5.1). For the results reported in this chapter, the model parameters are given in Table 5.1. In this implementation of the MACCEPA-VD, the equilibrium position and joint stiffness are controlled by two servomotors (Hitec HS-7950TH and Hitec HSR-5990TG, respectively). A DC motor (Maxon A-max 22/110125) is attached to the joint to serve as the damping motor, whose damping effect is controlled by the control unit introduced in §4.2.5. A current sensing module (Adafruit ina219

breakout) is connected in series with the electric load in the circuit to measure the high-side current and voltage to calculate the power of regeneration in real-time. A potentiometer (ALPS RDC503) is used to measure the joint angle. The velocity is then estimated by finite differences on the position data. The software architecture is based on the open-source Robot Operating System (ROS), where the control command is published to a ROS message, which is then subscribed to by a microcontroller (Arduino mega2560) to control the servomotors and the damping unit.

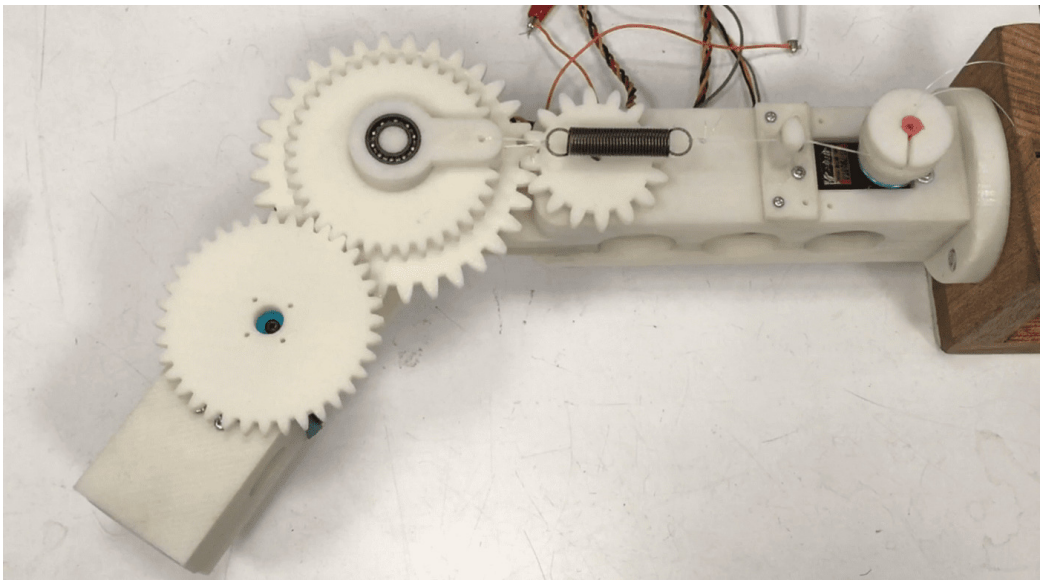


Figure 5.1: Physical implementation of MACCEPA Van Ham et al. (2007) with variable damping Radulescu et al. (2012).

Symbol	Value	Unit	Meaning
B	3.6	cm	length of lever arm
C	13.5	cm	pin displacement
r	1.5	cm	radius
κ	394	N/m	spring constant
m	0.0036	kgm^2	inertia
b	0.0077	$\mathrm{Nms/rad}$	friction coefficient
n_d	40		gear ratio

Table 5.1: MACCEPA-VD model parameters.

5.2.2 Control of the variable impedance robot

The variable impedance mechanism has intrinsic redundancy in its internal actuation. Optimal control has been demonstrated to be a straightforward and simple way to resolve this redundancy Braun et al. (2013) and efficient numerical solutions are available through local, iterative algorithms, such as ILQR. In the experiments reported here, ILQR is used to design the control sequence for the robot on the fly, as each reaching target is generated.

To determine the optimal control sequence, ILQR requires the robot dynamics formulated as a state space model.

Optimal control formulation

To represent the problem as an optimal control problem the dynamics are formulated as a state-space model $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u})$, where $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6)^\top = (q, \dot{q}, \theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2)^\top \in \mathbb{R}^6$ denotes the state space vector, $\mathbf{u} = (u_1, u_2, u_3)^\top \in \mathbb{R}^3$ is the control input, and $\mathbf{f}$ is defined as

$$\mathbf{f} = \begin{cases} x_2 \\ (\tau_s(x_1, x_2, x_3) - (d(u_3) + b)x_2)m^{-1} \\ x_5 \\ x_6 \\ \beta^2(u_1 - x_3) - 2\beta x_5 \\ \beta^2(u_2 - x_4) - 2\beta x_6 \end{cases} \tag{5.1}$$

The motor angles θ_1, θ_2 are controlled by $u_1 \in [-\pi/3, \pi/3], u_2 \in [0, \pi/3]$ respectively. The servomotor dynamics (3.29), (3.30) are assumed to behave as a critically damped system,[1] with $\beta = 25$. The damping coefficient $d(u_3)$ depends on control input u_3 and is calculated according to the damping scheme used (*i.e.,* (4.1), (4.2) or (4.7) in Chapter 4).

[1]Here, the coefficient β is chosen empirically to fit the step response of the servomotors.

The task of point-to-point reaching is captured by the cost function

$$J = \int_0^{t_f} (w_1(q - q^*)^2 + w_2 F_s^2 + w_3(u_3 - 0.5)^2 + w_4 \mathbf{u}^\top \mathbf{u})\, \mathrm{d}t \qquad (5.2)$$

where q^* is the reaching target, F_s is the spring force, and t_f is the reaching duration (for simplicity, in the experiments reported below, this is fixed at $t_f = 1.5$s for each movement). Here, the first term represents the reaching error and has weight $w_1 = 1000$ and the second term penalizes the squared spring force ($w_2 = 1$) which accounts for minimizing energy consumption.[2] The third term penalizes deviation of damping control from 0.5 (since this is known to be the point at which the regeneration coefficient is maximized, see §4.2) to encourage energy regeneration (in the experiments reported here it is weighted at $w_3 = 500$). The last term is added for regularization of the optimal control solution ($w_4 = 10^{-6}$). Note that, it is possible to use predicted regeneration power to replace the third term, however, it may result in behaviours sensitive to modelling errors of both dynamics and power regeneration.

5.3 Consecutive point-to-point reaching experiment

The task chosen to evaluate the proposed scheme is consecutive point-to-point reaching to random targets. The experimental procedure is as follows.

A list of $N = 25$ locations are generated sequentially as targets for reaching. Each target is drawn uniform randomly (*i.e.,* $q^* \sim U[-\pi/3, \pi/3]$), with the minimal distance between the target and the preceding one restricted to be at least $\pi/3$, to exclude very short-range movements. After generation, each target is fed to the cost function (5.2), and ILQR is used to determine the optimal control sequence for the movement under the dynamics (5.1), utilizing the whole control input space (*i.e.,* where all three control variables u_1, u_2 and u_3 are exploited to seek the optimum). The solutions are then executed on the plant, and the resultant joint trajectories and regenerated current are recorded. This procedure is

[2] A full model of the energy consumption of this actuator is not available. However, due to its mechanical design (with the spring pre-tension motor working against the spring), the overall electrical power consumption is monotonically related to squared spring force.

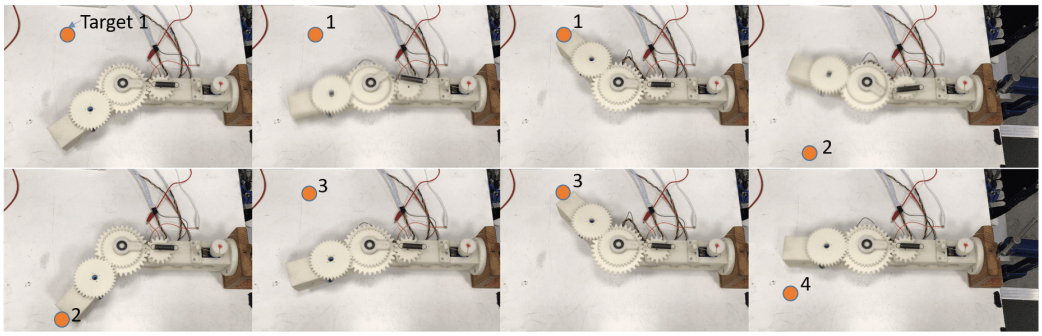


Figure 5.2: Snapshots of consecutive point-to-point reaching in the VSVD condition. The orange points overlaid show the reaching targets.

repeated $M = 20$ times to get a total of 500 recorded trajectories for performance evaluation. (In the below, this is termed the *variable stiffness variable damping* (VSVD) condition.) Figure 5.2 shows snapshots of reaching movements made to a typical sequence of targets.

For comparison, using the same reaching targets, the above procedure is repeated under three further conditions. Namely, these are:

1. *fixed stiffness and fixed damping* (FSFD): a baseline set where u_1, u_2 and u_3 are held at to constant values (in this case, reaching occurs by setting $u_1 = q^*$ prior to the onset of each reaching movement);
2. *fixed stiffness and variable damping* (FSVD): only the damping control u_3 is optimized, while u_1 and u_2 are held at fixed values;
3. *variable stiffness and fixed damping* (VSFD): the damping command u_3 is fixed and the equilibrium position and stiffness control inputs are optimized.

In the conditions where the stiffness is fixed (*i.e.,* FSFD and FSVD), u_2 is set to the minimal stiffness motor angle[3] $\pi/6$. In the conditions where the damping is fixed (*i.e.,* FSFD and VSFD), the damping command is set at $u_3 = 0.5$, corresponding to the point at which the regeneration coefficient reaches its maximum.

[3]The minimal stiffness motor angle is selected empirically to add sufficient pretension of the spring to provide good reaching accuracy around the zero joint position.

5.3.1 Performance metrics

In order to quantitatively compare the results, four metrics are employed to take into account of movement performance as well as energy consumption and regeneration:

Settling time The time when the plant settles down. For a given trajectory, this is defined as the smallest time t where both velocity and acceleration are within the vicinity of zero, *i.e.*, $|\dot{q}_t| < \epsilon_1$ and $|\ddot{q}_t| < \epsilon_2$. In our experiments, ϵ_1 was chosen to be approximately 1% of the maximum measured velocity and ϵ_2 was chosen to be 1.5% of maximum measured acceleration.

Overshoot The deviation of the joint position from the target point after overshooting the target. It is defined as the integration of $(q_t - q^*)^2$ from the time at which the target is first reached until the plant settles.

Energy consumption E_{in} computed by integrating P_{in} (defined in §2.3).

Regenerated energy E_{rege} computed by integrating the measured regeneration power.

Each of these are computed using the experimentally recorded data from the robot. For each trial, the settling time and overshoot of N trajectories are averaged and the energy regeneration and consumption are accumulated.

5.4 Results

The results for the four experimental conditions are reported in Table 5.2 and their

Experiment	Settling time (s)	Overshoot ($10^{-2}\text{rad}^2\text{s}$)	E_{in} (J)	E_{rege} (J)
FSFD	1.064 ± 0.039	4.980 ± 1.040	4.080 ± 0.552	0.152 ± 0.019
FSVD	0.923 ± 0.063	1.050 ± 0.350	3.989 ± 0.541	0.071 ± 0.014
VSFD	0.792 ± 0.061	0.650 ± 1.990	3.412 ± 0.358	0.089 ± 0.006
VSVD	0.780 ± 0.045	0.270 ± 0.120	3.067 ± 0.338	0.092 ± 0.010

Table 5.2: Performance metrics for the four experimental conditions computed on the recorded reaching data. Shown are mean $\pm$ standard deviation of each metric over 20 trials.

normalized scores are visualized in the radar chart in Figure 5.3, where the higher the score, the further out the line appears along that dimension. for example, the

energy regeneration score, denoted by γ_r, is the value of E_{rege} normalized according to its the maximum and minimum values in the four experimental conditions, and $\gamma_t, \gamma_o, \gamma_c$ are computed as the reciprocal of the normalized settling time, overshoot and energy consumption, respectively.

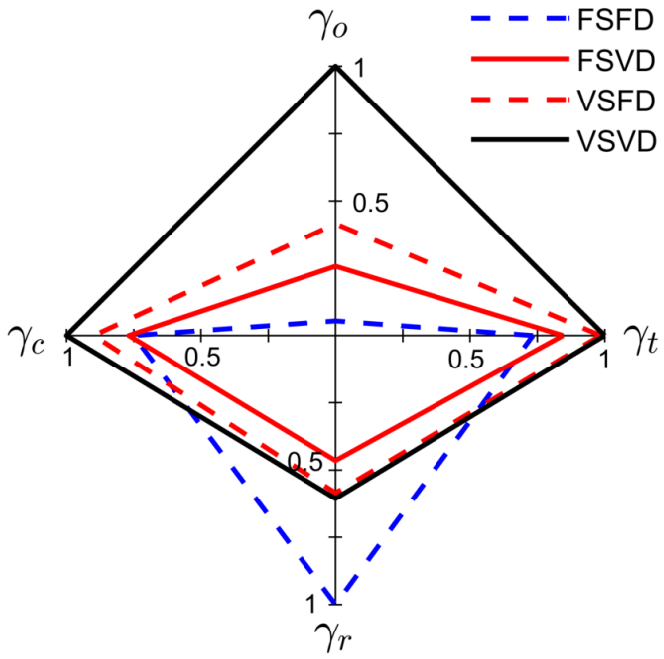


Figure 5.3: Radar chart showing the normalized reaching scores under the four experimental conditions. The higher the score the better the performance. Shown are mean scores over 20 trials.

Looking at each of these, it can be seen that in all conditions the damping module successfully regenerates power during the movement (see, for example, Figure 5.4(c) where a monotonic increase in accumulated energy is seen). Note, however, there is a discrepancy in the amount of energy regenerated and the corresponding performance in the reaching task.

As seen in Figure 5.3, the baseline condition (FSFD) harvests the most energy, by fixing the damping to the value that provides maximum regeneration coeffi-

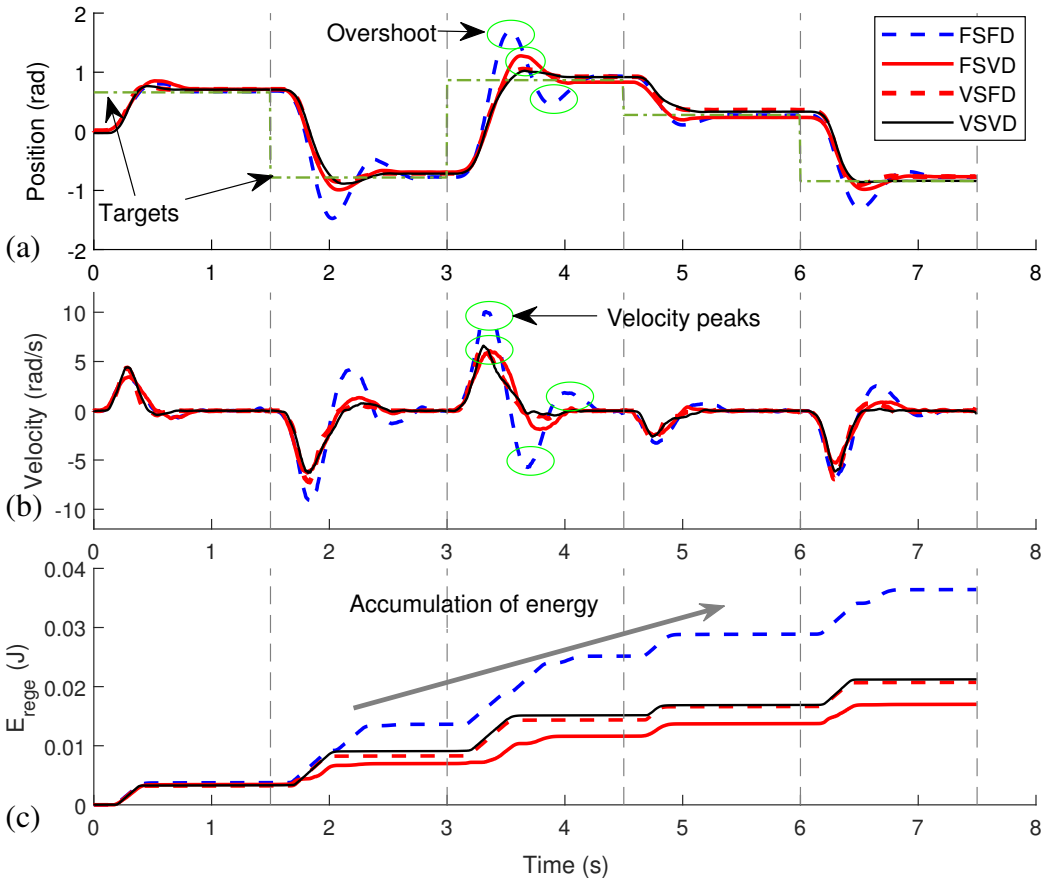


Figure 5.4: Recorded trajectories (a) joint position, (b) velocity, and (c) accumulation of regenerated energy E_{rege} of five typical examples of consecutive point-to-point reaching movements. The end of each movement is represented by the vertical dashed line. The dot-dashed line in (a) shows the target position for each phase of movement.

cient. Although this simple control strategy results in the most energy regeneration in the experiments, it sacrifices movement performance, scoring lowest in terms of the overshoot and settling time. Looking at the trajectory in Figure 5.4 (dashed blue line), it can be seen that there is significant overshoot for multiple targets. The enhanced energy regeneration also does not translate to lower energy consumption (see Figure 5.3 and Table 5.2). All this suggests that, although the stiffness and damping can be pre-tuned to give good performance for a specific movement, it can only be a solution for a specific task, and thus not suitable for a versatile actuator.

With FSVD, the movement performance in terms of overshoot and settling time is improved compared to FSFD, although the decrease in energy consumption is insignificant. The result confirms that variable damping can be utilized to improve the dynamic performance when the joint has fixed stiffness profile. It gives good overall dynamic performance for varied reaching targets. However, without exploiting variable stiffness, the variable damping cannot ensure energy efficient movements alone. In Figure 5.3 it can be seen that FSVD regenerates the least energy and there is no obvious improvement in terms of energy consumption.

By modulating the equilibrium position and stiffness VSFD performs moderately better on all performance metrics compared to FSVD. However, when all impedance variables are available to the controller, as in condition VSVD, it can be seen that the performance is significantly improved across the different metrics compared to the other conditions (see Figure 5.3). The energy efficiency is improved because there is less consumption (E_{in}) and more regeneration (E_{rege}). Additionally, although the average settling time is almost the same as that of VSFD, there is significantly lower overshoot.

Overall, these comparisons show that using variable damping in combination with an optimally exploited variable stiffness mechanism can contribute to both enhanced dynamic performance and improved energy efficiency (in terms of both consumption and regeneration).

5.4.1 Loss of regeneration through over-exploitation of damping

Comparing the performance of FSFD and FSVD it can be seen that a relatively modest improvement in reaching performance comes at the cost of a large reduction in the energy regeneration level. This seems surprising since FSVD also has available the possible strategy of keeping the damping fixed (although the damping can be varied, there is no imperative to do so). To better understand this behaviour, it is illuminating to examine in detail the control strategies chosen under the different conditions.

To examine this issue, the trajectories for reaching to a typical target position under the experimental conditions FSVD, VSFD and VSVD are plotted in Fig-

ure 5.5.[4] It can been seen that the plant controlled by the VSVD strategy hits the target position fastest. Both VSFD and VSVD controllers modulate the EP and stiffness motors to accelerate and decelerate the joint. However, only VSVD is allowed to adjust the damping command, which leads to an optimal strategy that lowers the damping level during the acceleration phase so that it reaches the target faster. The FSVD controller moves the EP position smoothly towards the target joint position. Without utilizing the "lead and lag" strategy, it has to apply a very high damping effect to brake the movement.

Looking at Figure 5.5(a), the accuracy is relatively good for each of the conditions (with significantly larger overshoot in the FSVD condition). However, to achieve this accuracy, the FSVD controller, being unable to modulate speed by any other means, modulates the damping over its full range, maintaining a high damping constant during braking ($0.2 < t < 0.5\,s$), see Figure 5.5(d). As the rate of energy regeneration has its maximum at $u_3 = 0.5$, the more time the damping is held away from its medium value, the lower the total accumulated energy will be (see §4.2). In contrast, the damping command of VSVD remains low during the acceleration phase to decrease the loss and required input power, but then maintains the damping command close to the maximum regeneration damping level for the remainder of the movement. This is thanks to its ability to modulate the joint stiffness and thereby effect braking by an alternative means (see Figure 5.5(b) and Figure 5.5(c)). The variable stiffness therefore provides some flexibility in control to prevent the *over-exploitation of variable damping* and the associated loss of regeneration.

5.5 Conclusion

In this chapter, we implemented the damping module developed in Chapter 4 into our 1-DOF VIA a MACCEPA type actuator. To investigate the use of variable regenerative damping for long-term operation, we designed a stochastic consecutive reaching task to examine the movement performance and energy efficiency. The experimental study shows that exploiting variable stiffness and variable damping

[4]The trajectory under each experimental condition was recorded by a single run.

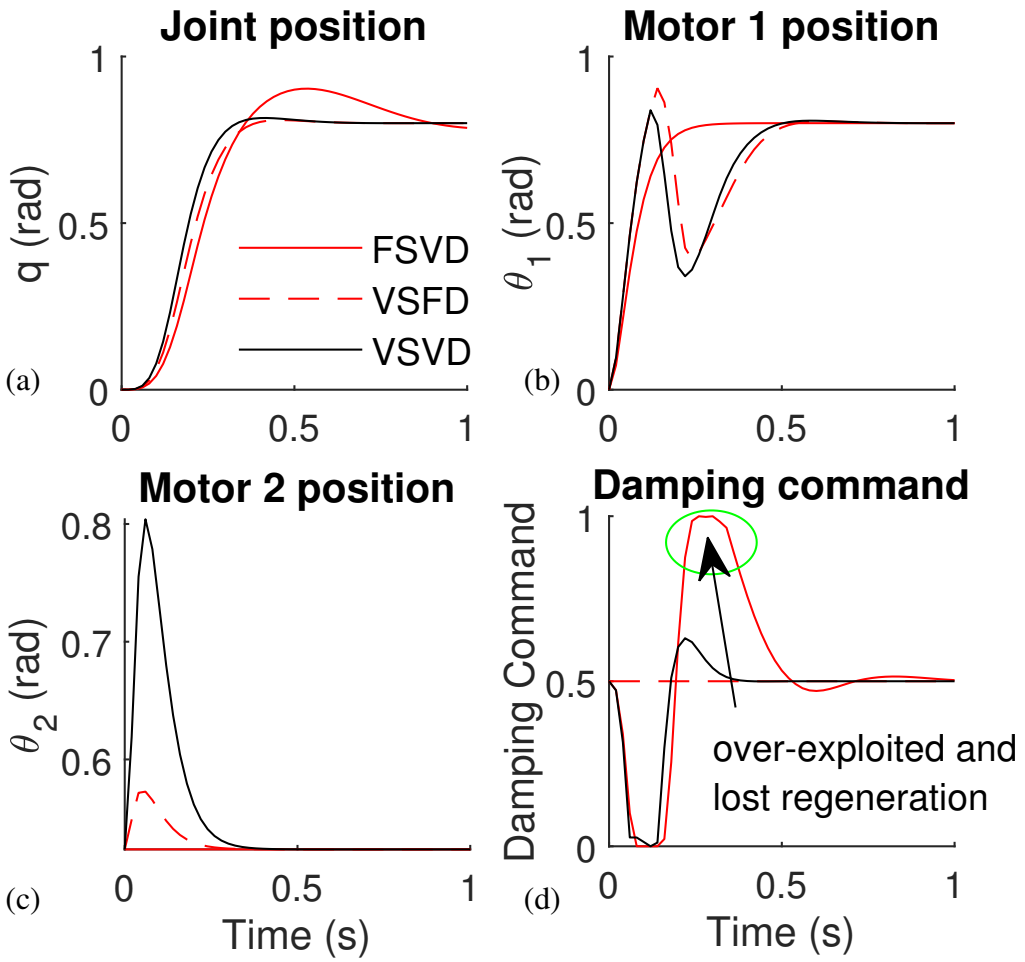


Figure 5.5: Comparison of a single reaching movement to the same target position under the conditions FSVD, VSFD and VSVD. The figure shows the (a) joint position, (b) motor 1, (c) motor 2, and (d) damping command against time.

is desired, in such a way that there is more flexibility to prevent over-exploitation of variable damping and loss of regeneration capability.

The tasks considered in this paper are movements in free space without external perturbations. It would be interesting to investigate the use of the damping module for more complex long-term behaviours in presence of external perturbations or unpredicted environments, such as pick-and-place different objects with unknown weights, long-distance locomotion using a bipedal platform.

A limitation of the experiment presented in this section is that the minimal stiffness pretension was set at a fixed value. The control cost term in the cost

function that penalizes deviation of the stiffness motor from this minimal pretension angle, results in the motor returning to the preset by default. This makes sense because sustaining pretension causes electrical consumption in the servomotor. It remains a question how to fully exploit the elasticity when the stiffness at transition should be optimally determined according to the subsequent movements.

We will continue to explore this problem by extending the optimal control to cope with a sequence of sub-movements.

Chapter 6

Exploiting Variable Impedance for Efficient Sequential Movements

In this chapter we consider sequential movements, which can be seen as an abstraction of the consecutive point-to-point reaching task used in the previous chapter. There, the possibility of planning a current movement according to its subsequent actions was neglected. Composite optimization of a long sequence in practise is often time-consuming and hard to scale up. Therefore, in this chapter we present some new thoughts, method and experiments to shed light on exploiting elasticity for energy efficient sequential movements.

6.1 Introduction

Intrinsically compliant robots typically have elastic components for stiffness modulation and such elements are capable of storing elastic energy. The field of robotic locomotion has seen a series of successful developments of energy efficient robots with elastic joints or springy legs that can exploit this energy storage. It is of great interest to apply the same principle to robotic manipulators such that soft robots can behave in a human-like energy efficient way for a wide variety of tasks.

Biological springs, like tendons and various elastic elements in muscles, are embedded in humans and animals and make them highly efficient runners and jumpers (Roberts (2016)). Utilizing elastic energy storage and recoil, which is

associated with optimizing muscular stiffness and transition timing, is a crucial skill that can be practised and improved for many other athletic activities, not limited to locomotion (Wilson and Flanagan (2008)).

Physical compliance incorporating elastic components is prominent for energy efficient lower limb locomotion (Reher et al. (2016); Roozing et al. (2016, 2019)). Also, they have been demonstrated to reproduce the skill of *energy buffering* in explosive movements such as throwing (Wolf and Hirzinger (2008); Braun et al. (2013)). Storing and discharging elastic energy, which was described as "skill of compliance" by Okada et al. (2002), can amplify the output power, exceeding the power limit of the drive motor. Other recent studies attempt to improve energy efficiency for cyclic manipulation tasks, *e.g.,* repetitive pick-and-place (Matsusaka et al. (2016)) and dribbling a basketball (Haddadin et al. (2018)).

However, many tasks in unstructured environments are not periodic and *variable physical impedance* is hard to fully exploit. For instance, the objects to be picked and placed may be located at random positions. The task given to a robot may consist of a sequence of different types of actions, such like "reach a cup, grasp it, and pour the water". These non-periodic but *sequential* tasks more commonly involve upper limbs and are complicated by their greater diversity. The problem of task-oriented sequential movement generation — in the context of compliant robotics — faces the difficulty imposed by inherent actuation redundancy. The control redundancy of the actuators, which is somehow equivalent to the muscle redundancy of musculoskeletal arms, makes it non-trivial to optimize the movements in the "muscle space".

Energetic economy is of great importance to reproduce human-like skilled movements. Researchers have embraced the notions of movement economy or efficiency since 1980s to understand and model human neuromuscular control of skilled movements (Nelson (1983); Sparrow and Newell (1998); Todorov and Jordan (2002)). The emergence and learning of complex motor skills can be explained as an optimization process aiming at minimizing metabolic energy expenditure subject to task, environment and organism constraints. Reduction of metabolic cost of human movements during training and practice has been verified by empirical studies (Lay et al. (2002); Huang et al. (2012)). However, for robots driven by VIAs (VIAs) which are viewed as the mechanical counterparts

of humans and animals, there lacks a systematic optimal control approach to optimize energy efficiency of complex skills modelled as sequential movements. To address this, our work postulates that such a framework should consider the following aspects:

1. **Cost function weighting**. When the form of the cost function is determined, the weighting parameter can be adjusted to tune energy efficiency. For simple quadratic control effort, the weight for each sub-movement need not be the same and can be optimized according to realistic energetics (by estimation or measurement).
2. **Variable impedance exploitation**. A movement can adjust physical impedance at the transition phase to improve its subsequent movements.
3. **Relative timing**. Temporal characteristics affect the energy efficiency. For a given time horizon, the relative timing is of importance for skilled efficient movements.

These three issues have been addressed separately in the literature. For example, inverse optimal control or inverse reinforcement learning is capable of learning the cost function from human demonstration (Mombaur et al. (2010); Berret et al. (2011); Levine and Koltun (2012)). Nakanishi et al. (2016) exploited variable stiffness actuation for *multiphase* movements by optimal control, where a brachiation task is used for demonstration. Nakanishi et al. (2011) extended optimal control (OC) to include optimization of movement durations. An analogue via approximate inference was provided in Rawlik et al. (2010). Other works focus on optimizing the sub-goals or attractors of movements encoded by dynamical systems (Toussaint et al. (2007); Stulp et al. (2012)).

However, rarely have existing approaches addressed the above targets in the sequential context within one framework. Also, many optimization-based methods rely on combining cost functions of subtasks into a composite one, which intensifies the cost function shaping issue when competing terms join together.

Therefore, this chapter proposes a hierarchical approach that is capable of optimizing the three aspects identified above and mitigate the cost function shaping issue (Wu and Howard (2020b)). More specifically, we employ a bi-level structure to encapsulate low-level OC for sub-movement generation into an outer loop of iterative policy improvement, thereby the benefits of both OC and RL are lever-

aged. The high-level optimization formulated as a reinforcement learning problem enables optimizing the trade-off balance concerning (low-level) (1) cost function weighting, (2) variable impedance exploitation and (3) transition timing for minimal *realistic* energetics. The associated high-level policy parameters can be optimized in a derivative-free fashion by a black-box optimization (BBO) method for policy improvement suggested by Stulp and Sigaud (2013). It can be viewed as a simplification of the RL algorithm PI^2 (Theodorou et al. (2010a)), which closely resembles an evolution strategy (ES) (μ, λ)-ES, the backbone of CMA-ES algorithm (Hansen and Ostermeier (2001)). At the low-level OC naturally resolves the actuation redundancy and exploit variable impedance of VIAs(Braun et al. (2012, 2013)), for which there exists efficient solvers *e.g.,* ILQR (Li and Todorov (2004); Tassa et al. (2014)).

The rest of this chapter is organized as follows. In §6.2 we discuss relevant literature and concepts. §6.3 first introduces a simple OC example of point-to-point reaching on the single joint VIA. By investigating the *efficient frontiers* of the OC problem we show how tuning the hyper-parameters[1] (of the OC problem) affects the energy efficiency trade-off and how the reinforcement learning problem is formulated. The proposed method is introduced in §6.4. Its effectiveness is evaluated by consecutive reaching tasks on a real VIA robot. Simulations demonstrate significant energy efficiency improvement and a reduction of electrical consumption of about 30% is recorded on the hardware. Conclusions and future works are covered in §6.6.

6.2 Related Work

6.2.1 Sequential movements

Sequential movements are commonly found in human daily life, from jaw movement for speech, finger movement for playing musical instruments, to many athletic whole body actions. How can these skilful human movements be learnt, executed and improved? Central to that is whether a hierarchical structure of

[1] In this chapter it assumes that the term "hyper-parameter" is associated with Optimal Control problems unless otherwise specified.

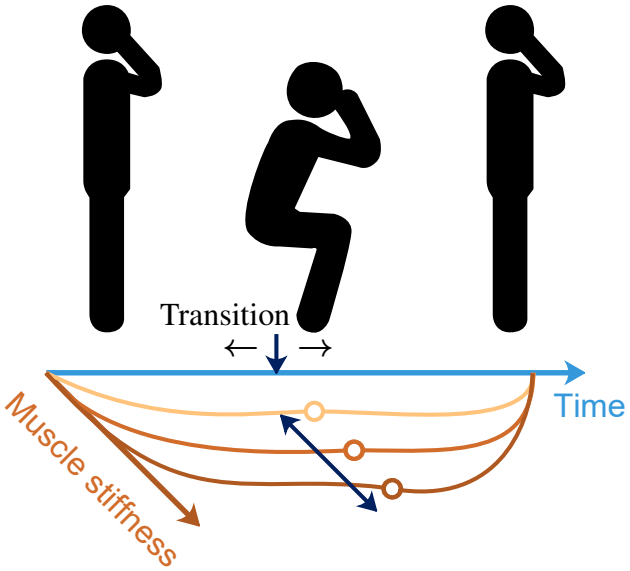


Figure 6.1: Human can acquire new skilled movement by sequencing simpler motion primitives. A squat can be composed of crouching and rising-up, and the corresponding variables in the sequential context can be improved through practice. Possible ways to optimize the squat towards higher energy efficiency are: (i) adjust transition timing, and (ii) modulate muscular stiffness.

representation, learning and control of movement sequences exists in the human brain. The hypothesis of hierarchical organization of movement planning was proposed a long time ago in the mid twentieth century by behaviourist Karl Lashley (Lashley (1951)). Recent experimental studies have provided evidence of hierarchical representation of movement sequences in the brain. For instance, Yokoi and Diedrichsen (2019) found that individual finger presses are represented in the primary motor cortex, whereas activities about the sequential context happen mainly in the premotor and parietal cortices.

In the robotics literature, sequential composition of controllers was employed by Burridge et al. (1999) to achieve dynamically dexterous robot behaviours. In robot learning control, motion generation of complex skills is often investigated at the task level and treated in a hierarchical manner. A complex skill can be learnt from human demonstrations by motion segmentation into movement primitives (Lucia et al. (2013)). Then a skilful movement can be composed by a "repertoire" (Schaal and Atkeson (2010)) of such sequenced submovements. By doing so it

is expected to realize more general motion intelligence and make robots master interactive tasks and tool use, which is a hallmark of human behaviour (Hogan and Sternad (2012)).

Humans can acquire a new skilled movement by sequencing simpler motion primitives and improving via practice. Although each individual movement is possible to be fine-tuned during training, the increased performance through practice can be clearly attributed to improvements in planning processes, as shown by Ariani and Diedrichsen (2019). Consider a squat (Figure 6.1) that can be composed of crouching and rising-up. By intuition, the contextual variables at the sequence planning level can possibly be transition timing, muscular stiffness, torque distribution, *etc.* Motegi and Matsui (2011) used OC to find optimal transition timing that can reproduce experimentally measured human squat movements. The role of stiffness was investigated by Bobbert (2001) also through biomechanical modelling and OC, which signifies the importance of *exploiting elastic energy storage*.

6.2.2 Optimization of sequential movements

Improvement of a sequential movement necessitates existence of redundancy in either representational level or control level. In the above squat example, the transition timing is not predefined by the task or sub-movements, and thus can be tuned. While for playing a piece of music, the tempo and rhythm are determined, then the transition timing is specified by the task objective and cannot be exploited.

It is easy to notice that sequentially combining the sub-movements, which are optimized with respect to their sub-goals, does not necessarily result in the optimal movement for the whole task. Look at the example illustrated in Figure 6.2, the minimal-jerk trajectory[2] of a via-point task from point A to C via B (at a specific time) results in a curved path in the X-Y plane, while the minimal-jerk model of a single point-to-point movement always shows a straight path. Consequently, if we sequence AB and BC (both individually are minimal-jerk) directly, the result-

[2]The analytical formulas to calculate the minimal jerk trajectory of point-to-point reaching and via-point reaching are given in Flash and Hogan (1985).

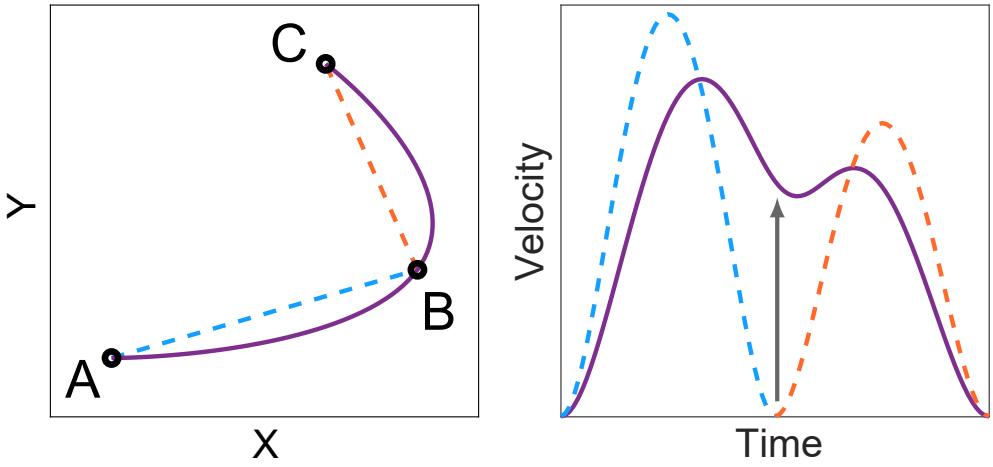


Figure 6.2: Minimal jerk trajectories AB, BC and a via-point movement AC. AB, BC are both individually minimal jerk trajectories, but simply sequencing them is not optimal for A via B to C. The optimal via-point minimal jerk trajectory is curved around B in the X-Y plane (right).

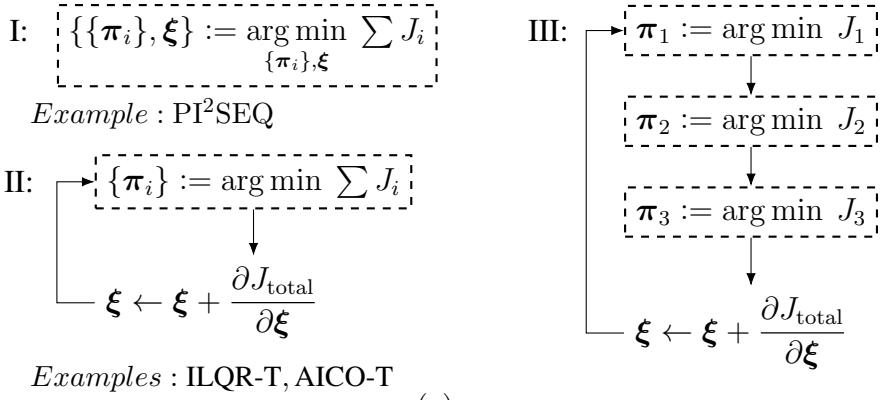
ing trajectory is not optimal w.r.t. the whole movement. The difference is simply due to that the velocity at the via point is not constrained to be zero. Though it is obvious, this common phenomenon in the kinematic domain shows an example of exploiting the *redundancy* of velocity profile when concatenating discrete movements.

Based on the above reasoning, it follows that when a sequence is generated by chaining movement primitives, it may be suboptimal without appropriately planning each individual considering the whole trajectory or its subsequent ones. In general, one can structure the problem as a composite optimization or tackle it hierarchically. Depending on whether either way is adopted, or both, there are three possible approaches, as depicted in Figure 6.3. Throughout this chapter, π denotes the control policy, $\boldsymbol{\xi}$ represents the vector of policy parameters to be optimized via reinforcement learning, J is used for cost function.

To narrow down our discussion, we mainly consider the applications for (i) optimization of cost function weighting, (ii) exploitation of variable impedance, and (iii) optimizing temporal parameters such as the time horizon and relative timing.

To avoid confusion, the sequential movements/tasks considered in this chapter are sequences in a predefined order. The problem of planning the order of ex-

ecuting a set of actions for a given task is not within the scope of this work. This kind of task planning problem does not predefine an order of executing subtasks. Therefore it needs a higher level planning to figure out the best order to chain the submovements, typically from a discrete set of actions (Manschitz et al. (2015)).



(a)

	I	II	III
Optimize cost function weighting	No	Yes	Yes
Exploit variable impedance	Yes	Yes	Yes
optimize temporal parameters	Yes	Yes	Yes
Avoid redesign of composite cost function	No	No	Yes

(b)

Figure 6.3: (a) Three plausible types of approaches for sequential movement optimization. Dashed rectangle means a full optimization loop. (b) The table summarizes comparison of type I-III. For clarity, in Type III only 3 sub-problems are shown to visualize a sequence.

Composite optimization

Composite optimization here means optimizing w.r.t. a composite cost function that consists of the objectives of subtasks. For instance, an optimization-based approach usually considers the via-point problem by defining the cost function as

$$J = (\mathbf{x}(t_v) - \mathbf{x}_v^*)^\mathsf{T} \mathbf{H}_v (\mathbf{x} - \mathbf{x}_v^*) + (\mathbf{x}(t_f) - \mathbf{x}_f^*)^\mathsf{T} \mathbf{H}_f (\mathbf{x} - \mathbf{x}_f^*) \tag{6.1}$$

Here $\mathbf{x}$ is the state vector of the problem, $\mathbf{x}_v^*, \mathbf{x}_f^*$ are the via-point and final targets respectively, $\mathbf{H}_v, \mathbf{H}_f$ are diagonal matrices to penalize the deviation, and t_v, t_f

represent the fixed via-point time and final time. The optimal control $\mathbf{u}(\mathbf{x}, t) = \boldsymbol{\pi}(\mathbf{x}, t)$ with corresponding policy $\boldsymbol{\pi}$ is the one that minimizes the cost functional. The above minimal jerk via-point problem is one example that has analytical solution (Flash and Hogan (1985)). The shortcoming of this is that if t_v is allowed to be adjusted, the optimization of (6.1) with a guess about t_v may leads to suboptimal solutions.

Let us first consider the possibility to simultaneously optimize the control sequence and the hyper-parameters like t_v in the above example. This is categorized as Type I in Figure 6.3. For many non-linear real problems arising in robotics, a classical method is to convert the OC problem into a non-linear programming problem. Considering the computational efficiency, a more efficient paradigm for learning control is to transform the representation of the control policies into a lower-dimensional space, and then optimize the policies and their hyper-parameters simultaneously. For example, Stulp et al. (2012) implemented the (model-free) reinforcement learning algorithm PI^2 for sequential tasks (termed as $\mathrm{PI}^2\mathrm{SEQ}$), with the help of dynamic movement primitives (DMP) for trajectory encoding using dynamical systems. The shape parameter of trajectories and the attractors of dynamical systems are optimized together, so that the trajectory and its final state is optimized for all subsequent actions. The limitation of composite cost function is that it faces the cost function shaping issue. When competing terms from different subtasks come together, optimality of sub-movements may be no longer achievable. In order to achieve optimality for all subtasks the formulation of the cost functions have to be redesigned.

Hierarchical optimization

The second possible approach is to construct the optimization problem hierarchically. As shown by Type II in Figure 6.3, it is hierarchical in the sense that an inner loop and an outer loop optimize the control policies and hyper-parameters separately. Various previous studies addressing the multiphase optimal control can be found in this type. To name a few, temporal optimization with ILQR (ILQR-T) and approximate inference (AICO-T) was proposed by Nakanishi et al. (2011) and Rawlik et al. (2010) respectively. Nakanishi et al. (2011) used finite difference to

compute the gradient of total cost w.r.t. change of time durations. The evaluation of the gradient is based on running the time-scaled augmented control and hence is very efficient. This is done by leveraging a technique that maps the real time to a canonical time. It was demonstrated by Rawlik et al. (2010) with a similar technique on a via-point task, where the algorithm finds an optimal relative timing. In case J_{total} is non-differentiable w.r.t. $\boldsymbol{\xi}$, one can utilize derivative-free methods (Conn et al. (2009)) such as trust region technique (Yuan (2015)) and evolutionary strategy (Hansen and Ostermeier (2001)) in the outer loop.

This hierarchical structure coincides with the so-called "bi-level" problem in inverse optimal control (Mombaur et al. (2010)). In inverse optimal control the outer loop optimizes the cost function shaping to match data demonstrated from a human. Of interest here is the fact that the objective J_{total} in the outer loop need not be the same as the composite cost. Suppose that, for the speed and robustness of optimization, the subtasks may be described with simple quadratic terms such as traditional "control effort", or even have different energetic functions individually, but on the high level, the parameter can be updated according to more realistic cost estimator or physical measurement. This potential can be realized within the bi-level architecture.

Note that, since Type II also employs a composite cost function in the inner loop, it shares the same shortcoming with Type I that composite optimization may fail to achieve optimality for all subtasks and thus need redesign. To overcome the drawback, we propose to optimize the sub-movements according to their own cost function as well as integrate the hierarchical (bi-level) architecture, which leads to Type III (Figure 6.3). The comparison against previous two types is summarized in the table (Figure 6.3(b)).

6.3 Problem Formulation

In this section we first present a OC model of a single joint driven by a VIA, followed by an investigation of energy efficiency based on the concept of efficient frontiers. The intuition gained thereby helps with justifying the problem formulation. Finally, a reinforcement learning problem is formulated that enables optimizing high-level parameters using policy improvement methods.

6.3.1 A simple reaching movement model

Consider a point-to-point fast reaching task using a single-link robot driven by a VIA. We choose MACCEPA-VD which was developed and used in previous chapters. The system modelling is given in §3.2. The fast reaching task is represented by a cost functional

$$J(x(\cdot), u(\cdot)) = H(x(t_f)) + \int_0^{t_f} l(x(t), u(t), t)\,\mathrm{d}t \tag{6.2}$$

$$H(x(t_f)) = 1000(q(t_f) - q^*)^2 \tag{6.3}$$

$$\begin{aligned} l(x(t), u(t), t) = {} & 1000(q(t) - q^*)^2 \\ & + w_e((u_1(t) - q^*)^2 + u_2^2(t) + 10^{-3}(u_3(t) - 0.5)) \end{aligned} \tag{6.4}$$

An optimal control problem can be formulated as to seek an optimal control $\mathbf{u}(t) \in U \in \mathbb{R}^3$ constrained by its admissible set $U = \{\mathbf{u} \in \mathbb{R}^3 \,|\, \mathbf{u}_{\min} \preceq \mathbf{u} \preceq \mathbf{u}_{\max}\}$, that minimizes the cost functional (6.2) and subject to the SSM of the robot dynamics. In the cost function, w_e serves as a weighting parameter to enable adjustment of the performance-cost trade-off.

In addition to trade-off balance via cost function weighting, the stiffness at transition could have a significant influence on the energy efficiency of the subsequent movement. This is explained as follows.

6.3.2 Efficient frontiers of optimal control

The Efficient Frontier (EF) is a common tool to examine the trade-off of two competing objectives in an optimization problem. In this work the problem can be interpreted as optimizing the task performance while minimizing the energy cost. The EF is then the set of optimal solutions that achieve the best performance at a defined energy cost. The above OC problem has an efficient frontier by varying the weighting parameter w_e. Then the distribution of optimal solutions can be visualized in performance-cost plane.

In addition, to investigate how pre-stored elastic energy affects the energy efficiency, we generate the optimal solutions by ILQR for different values of w_e, with a certain minimal spring pretension to produce an EF. Multiple EFs are generated

by changing the condition of minimal spring pretension. This is done by setting the initial stiffness motor angle $\theta_2(0)$ and the lower bound $u_{\min}^{(2)}$ of u_2 to a preset value p_s, *i.e.*, let $\theta_2(0) = u_2^{\min} = p_s \in P_s := \{p_s \in \mathbb{R} \mid \theta_2^{\min} \leq p_s \leq \theta_2^{\max}\}$.

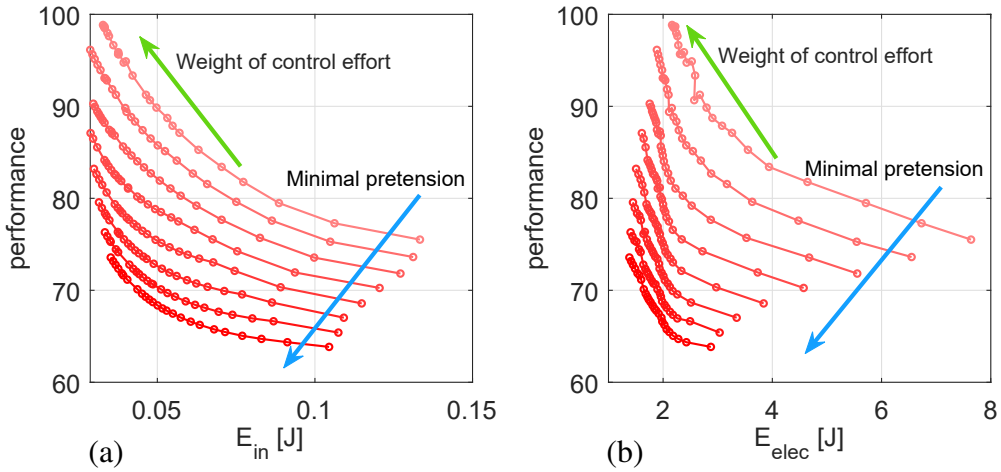


Figure 6.4: Efficient frontiers : (a) fast reaching performance against input mechanical work, and (b) fast reaching performance against electrical work. Each efficient frontier shows the optimal control solutions by varying the weight w_e of control effort term (shown by the green arrow) in the cost function, with a certain minimal spring pretension. The green arrow indicates the direction of increasing the weight. The spring preset parameter p_s is adjusted by the servo M_2 from $0.1\,\mathrm{rad}$ to $1.5\,\mathrm{rad}$ with increment of $0.2\,\mathrm{rad}$. Increasing the minimal spring pretension (shown by the blue arrow) moves the efficient frontier downward, which means an increased overall energy efficiency.

The results are shown in Figure 6.4. The vertical axis represents the reaching accuracy performance, which is the terminal cost (6.3) plus the integral of the first term of running cost (6.4). The horizontal axis is the energy cost, measured by positive input mechanical work[3] E_{in} and electric work E_{elec}, both estimated by

[3]We assume that the motors are not back-drivable, thus no negative mechanical work to the motors can be regenerated. Similarly, the electrical energy is defined as the integral of the positive part.

simulation.[4]

$$E_{\text{in}} = \int [P_{\text{in1}}]^+ + [P_{\text{in2}}]^+ \, \mathrm{d}t \tag{6.5}$$

$$E_{\text{elec}} = \int [P_{\text{elec1}}]^+ + [P_{\text{elec2}}]^+ \, \mathrm{d}t \tag{6.6}$$

where $[\cdot]^+ = \max(0, \cdot)$. For in-depth analysis of modelling motor energy consumption, we refer the readers to Verstraten et al. (2016). Calculation of the mechanical and electrical power is given in Appendix B.

Looking at Figure 6.4, when increasing the weight of control effort w_e (as shown by the direction of green arrow), both mechanical and electrical consumption are decreased, with some loss of reaching performance. It demonstrates that even with a simple quadratic control cost, it is still possible to tune the trade-off between performance and realistic energy measures. Moreover, it can be seen that, by increasing the minimal spring pre-tension p_s, the efficient frontiers move towards the bottom-left, which signals an overall improvement of energy efficiency.

Overall, the above investigation based on the tool of efficient frontier suggests that control cost weight w_e and minimal stiffness p_s can be treated as hyper-parameters that tune the performance-cost trade-off according to realistic energy measures.

6.3.3 Reinforcement learning formulation

Based on the previous rapid reaching OC model, let us now consider a consecutive reaching task for example, that requires the arm to reach a sequence of targets $\{q_i^*\}_{i=1}^{N_s}$ from initial state $\mathbf{x}_0$, where N_s is the number of subtasks. The sequential movement $\mathcal{S} \coloneqq \{\mathcal{M}_i\}_{i=1}^{N_s}$ consists of N_s sub-movements generated by solving optimal control problem (OCP). The sub-problems is denoted as $\{\mathrm{OCP}_i\}_{i=1}^{N_s}$. We define $\boldsymbol{\xi} \in \{\boldsymbol{\xi} \in \mathbb{R}^{d_p} \mid \boldsymbol{\xi}_{\min} \preceq \boldsymbol{\xi} \preceq \boldsymbol{\xi}_{\max}\}$ to be the stacked vector of weighting parameter $\mathbf{w}_e = \{w_e^{(i)}\}$, stiffness parameter $\mathbf{p}_s = \{p_s^{(i)}\}$, and movement durations

[4]Note that, the accuracy of estimating $E_{\text{in}}, E_{\text{elec}}$ is very sensitive to simulation step size. For ILQR we typically use time step $\Delta t = 0.02$. While computing $E_{\text{in}}, E_{\text{elec}}$ is based on simulation (of forward dynamics) with $\Delta t = 0.001$.

$\mathbf{t}_d = \{t_d^{(i)}\}$.[5] $\boldsymbol{\xi}_{\min}, \boldsymbol{\xi}_{\max}$ are lower and upper bound of $\boldsymbol{\xi}$. Note that, depending on the type of task at hand, $\mathbf{p}_s$ may have different meaning. For example, as in §6.3.2 it is used to set the minimal stiffness motor command. By doing so it constrains the minimal elastic energy to be stored and sets a target for the motor.

Our problem is to find an energy optimal trajectory $\mathcal{S}$ and $\boldsymbol{\xi}$ that minimizes energy cost while achieving all sub-goals. Mathematically, it is formulated as to minimize the episodic cost:

$$J(\mathcal{S}) = J_e + \mathcal{C} \cdot \max\{0, J_p - \bar{J}_p\} \tag{6.7}$$

The cost objective (6.7) is formulated as an *episodic* cost. J_e is the energy consumption, and J_p is the cost associated with task achievement. The amount of J_p exceeding an upper bound $\bar{J}_p$ is penalized by a large constant $\mathcal{C}$. The energy consumption can be estimated by a cost functional or measured on hardware. $\bar{J}_p$ is evaluated by solving $\{\mathrm{OCP}_i\}$ with initial $\boldsymbol{\xi}^{(0)}$.

6.4 Policy Improvement for Sequential movements

The policy improvement optimizes J in an iterative process. Figure 6.5 outlines the paradigm of our proposed policy improvement method which encapsulates OCPs at the low-level. It consists of the main steps of general policy improvement procedures: exploration, evaluation, and policy update. Different from the vanilla reinforcement learning from exploration and evaluation we have an inner loop to solve $\{\mathrm{OCP}_i\}$ sequentially.

The first step is to evaluate the initial trajectory with $\boldsymbol{\xi}^{(0)}$ given by the user. Once $\{\mathrm{OCP}_i\}$ are specified, we run ILQR to generate $\mathcal{S}^{(0)} = \{\mathcal{M}_i\}_{i=1}^{N_s}$ and obtain corresponding costs $J_e^{(0)}, J_p^{(0)}$.[6] The task performance constraint is set up by multiplying a tolerance factor $\sigma_{\mathrm{tol}} \in \{0 \cup \mathbb{R}^+\}$ with $J_p^{(0)}$, *i.e.,* $\bar{J}_p = (1 + \sigma_{\mathrm{tol}})\, J_p^{(0)}$. The tolerance factor is introduced for the user to trade-off the energy efficiency flexibly. A positive value allows the exploration for some samples that have worse performance so that the information may contribute to faster and more robust up-

[5] By convention, all vector quantities are assumed to be column vectors.

[6] For the initial trajectory, it is obvious that $J^{(0)} = J_e^{(0)}$

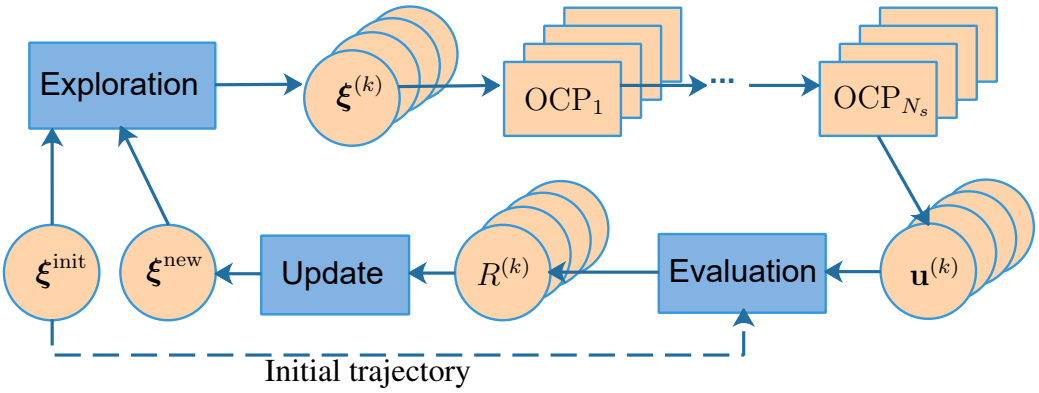


Figure 6.5: Diagram of policy improvement method.

dating towards the minimal energy cost.

6.4.1 Exploration and evaluation

The exploration phase generates K unconstrained perturbations in policy parameter space for K roll-outs. The perturbations $\tilde{\epsilon}_k \sim \mathcal{N}(\mathbf{0}, \gamma^{n-1}\boldsymbol{\Sigma}_\epsilon), (k = 1, ..., K$ is assumed to obey normal distribution, where $\boldsymbol{\Sigma}_\epsilon$ is the covariance matrix and $\gamma \in (0, 1)$ is the decay factor. Then the box constraint $\underline{\mathbf{b}}$ and $\bar{\mathbf{b}}$ is applied to yield

$$\boldsymbol{\epsilon}_k = \min(\max(\tilde{\boldsymbol{\epsilon}}_k + \boldsymbol{\xi}^{(n)}, \boldsymbol{\xi}_{\min}), \boldsymbol{\xi}_{\max}) - \boldsymbol{\xi}^{(n)} \tag{6.8}$$

$$\boldsymbol{\xi}^{(n)[k]} = \boldsymbol{\xi}^{(n)} + \boldsymbol{\epsilon}_k \tag{6.9}$$

When running each k-th roll-out, $\boldsymbol{\xi}^{(n)[k]}$ is used to specify sub-problems with $\mathbf{w}_e$ for cost functional J_i, $\mathbf{p}_s$ for stiffness motor constraint, and $\mathbf{t}_d$ for time horizon. Without loss of generality, we assume the time horizon $[t_0, t_f]$ of i-th sub-problem is from $t_0 = 0$ to $t_f = t_d^{(i)}$.

With these details, $\{\mathrm{OCP}_i\}$ are solved by ILQR sequentially to generate the sub-movements $\mathcal{M}_i$. The final state of $\mathcal{M}_i = \{\mathbf{x}_i, \mathbf{u}_i\}$ is taken as the initial state of its sequent problem, *i.e.*, $\mathbf{x}_{i+1}(0) = \mathbf{x}_i(t_f^{(0)})$. The energy consumption $J_e^{[k]}$ and task performance $J_p^{[k]}$ along the trajectory is then evaluated by running a forward pass of dynamics and control $\mathbf{u} = \{\mathbf{u}_i\}$. After running K roll-outs and collecting relevant costs, the total costs $J^{[k]}$ are calculated by (6.7).

6.4.2 High-level policy update

The policy update step (6.10)-(6.12) utilizes the reward-weighted averaging rule as introduced by Stulp and Sigaud (2013).

$$\tilde{J}^{[k]} = \frac{J^{[k]} - \min(\{J^{[k]}\})}{\max(\{J^{[k]}\}) - \min(\{J^{[k]}\})} \tag{6.10}$$

$$P_k = \frac{\exp(-c\tilde{J}^{[k]})}{\sum_{i=1}^{K} \exp(-c\tilde{J}^{[i]})} \tag{6.11}$$

$$\boldsymbol{\xi} \leftarrow \boldsymbol{\xi} + \sum_{k=1}^{K} P_k \boldsymbol{\epsilon}_k \tag{6.12}$$

First the cost $J^{[k]}$ is normalized according to their maximum and minimum by (6.10). The normalized cost $\tilde{J}^{[k]}$ is used to calculate probability P_k for k-th roll-out according to (6.11), where $c > 0$ is a constant.[7] Finally, the update is computed by the weighted averaging rule (6.12).

The above weighted averaging technique is simplified from PI^2 (Stulp and Sigaud (2013)) and converts the policy improvement method into a black-box optimization (BBO) method that resembles the evolutionary strategy (μ, λ)-ES, which is the basic form of CMA-ES algorithm (Hansen and Ostermeier (2001)). It is appealing because it can solve non-linear non-convex BBO problems with reasonable efficiency. Unlike CMA-ES, it does not have the covariance matrix adaption step. Instead, we manually specify a decay factor γ to gradually decrease the variance of perturbations.

A (μ, λ)-ES method consists of three steps: mutation, selection and recombination. The exploration phase corresponds to the mutation step. Then all samples are selected for policy update (recombination). The policy update step can be viewed as recombination of samples. In contrast to the reinforcement learning algorithm PI^2 that leverages the problem structure, ES treats the policy improvement as a BBO problem. Since the high-level optimization is solved as a BBO problem in our proposed policy improvement method, we label the high-level part of the whole method as a evolutionary strategy.

For better robustness of convergence, another technique employed is sample

[7] In our implementation we choose $c = 10$.

Algorithm 1 Optimization of sequential movements using OC-ES

Given: $\{J_i\}$, $\{\mathrm{OCP}_i\}$, $\boldsymbol{\xi}_{\min}, \boldsymbol{\xi}_{\max}$
Initialization: $\boldsymbol{\xi}^{(0)}, \gamma, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}, \mu, \sigma_{\mathrm{tol}}$
Generate $\mathcal{S}^{(0)}$ by solving $\{\mathrm{OCP}_i\}$, compute $\bar{J}$
repeat
  for $k = 1$ to K **do** ▷ k-th rollout
    Sample $\tilde{\boldsymbol{\epsilon}}_k \sim \mathcal{N}(\mathbf{0}, \gamma^{n-1}\boldsymbol{\Sigma}_{\epsilon})$ ▷ Unconstrained perturbations
    $\boldsymbol{\epsilon}_k = \min(\max(\tilde{\boldsymbol{\epsilon}}_k + \boldsymbol{\xi}^{(n)}, \boldsymbol{\xi}_{\min}), \boldsymbol{\xi}_{\max}) - \boldsymbol{\xi}^{(n)}$ ▷ Constrained perturbations
    $\boldsymbol{\xi}^{(n)[k]} \leftarrow \boldsymbol{\xi}^{(n)} + \boldsymbol{\epsilon}_k$
    Specify hyper-parameters and constraints of $\{\mathrm{OCP}_i\}$ according to $\boldsymbol{\xi}^{(n)[k]}$
    for $i = 1$ to N_s **do**
      $t_0 = 0, t_f = t_d^{(i)}, \mathbf{x}_i(0) = \mathbf{x}_{i-1}(t_f)$,
      solve $\{\mathrm{OCP}_i\}$, $\mathbf{u}_i = \arg\min J_i$
      $\mathcal{M}_i = \{\mathbf{x}_i, \mathbf{u}_i\}$
    end for
    Estimate $J_p^{[k]}, J_e^{[k]}$,
  end for
  Retrieve stored samples and append to dataset $\{J^{[k]}, \boldsymbol{\epsilon}_k\}$, $K' = K + \mu$
  Compute and normalize cost $\{J^{[k]}\}_{k=1}^{K'}$ by (6.7) (6.10)
  Update $\xi^{(n+1)}$ using (6.11) and (6.12)
  Keep μ best samples for sample reuse
until $\boldsymbol{\xi}$ converges or maximum number of iterations reached

reuse. After every update, we keep μ best samples among K roll-outs (at current iteration) for next update. Therefore, after the first iteration, we have $K + \mu$ samples. The exploration, evaluation and policy update procedures are repeated until $\boldsymbol{\xi}$ converges or reaches maximum steps. The whole algorithm is summarized in Algorithm 1 and termed as OC-ES, which stands for Optimal Control (at low-level) with Evolutionary Strategy (at high-level).

6.5 Applications and Evaluations

6.5.1 Energy efficient consecutive fast reaching

Task 1 - consecutive fast reaching

To evaluate our proposed method, a consecutive fast reaching task is designed to test on the MACCEPA-VD robot. The task requires the joint actuated by MACCEPA-VD to reach a sequence of three targets $\{q_i^*\}_{i=1}^3 := \{0.7, -0.35, 0.3\}$ (radians) rapidly within a fixed time horizon $T_i = 1$, for $i = 1, 2, 3$, from initial state $\mathbf{x}_0 = (0, 0, 0, \pi/24, 0, 0)^\mathsf{T}$. The cost functional J_i for each subtask is defined by (6.2) - (6.4). A single fast reaching problem was used by Radulescu et al. (2012) to investigate the role of variable damping for VIAs when an appropriate amount of damping is needed to suppress oscillation of movements. In our previous work in Chapter 5, a consecutive (stochastic) fast reaching experiment was conducted to examine how combination of variable stiffness and regenerative damping influence the overall energy efficiency. There, the minimal spring pretension between two movements was manually set by user experience, whereby the variable stiffness was not fully exploited in the sequential context.

For comparison, a benchmark is generated by using ILQR to solve the subproblems sequentially. The spring preset $p_s^{(i)} = \pi/24\,\mathrm{rad}$ is the lower bound of stiffness motor position command and weighting parameter $w_e^{(i)} = 1$ for $i = 1, 2, 3$. The resulting (approximately) optimal trajectory $\mathcal{S}$ is denoted by ILQR-0, and used as initial trajectory later for our proposed method.

6.5.2 Task 1: policy improvement with parametrized trajectory

The competing terms in the composite cost function may hinder the fulfilment of all sub-goals. To investigate this issue we directly optimize the trajectory and

stiffness profile simultaneously w.r.t. the composite cost function of **Task 1**

$$J(\mathcal{S}) = \sum_{i=1}^{3} J_i(\mathcal{M}_i) \tag{6.13}$$

$$J_i = 1000\left((q(t_f) - q_i^*)^2 + \dot{q}^2(t_f)\right) + \int_{t_0}^{t_f} 1000\,(q(t) - q^*)^2 dt \tag{6.14}$$

$$+ \int_{t_0}^{t_f} \left(100\,(\theta_1 - q_i^*)^2 + 100\,\theta_2^2 + 10^{-3}\,(\theta_3 - 0.5)\right) \mathrm{d}t \tag{6.15}$$

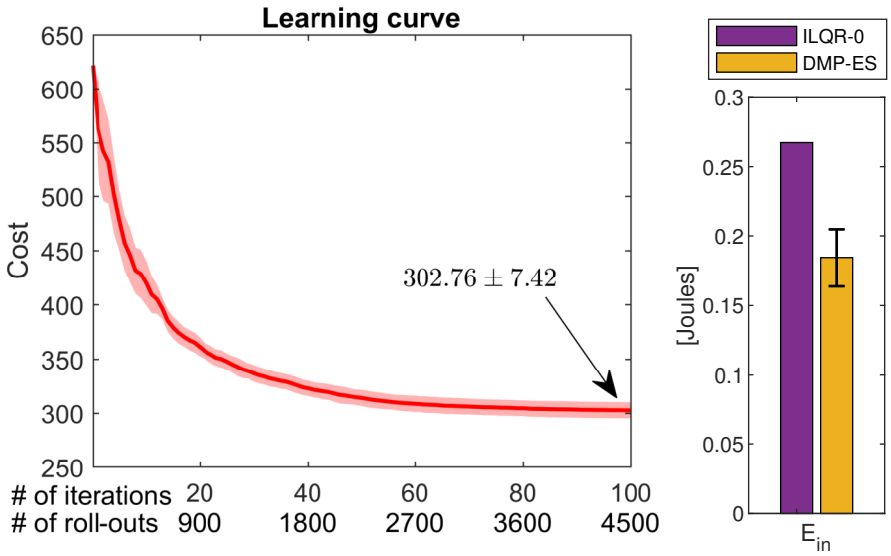


Figure 6.6: Learning curve (**left**) of PI^{2}SEQ for the consecutive fast reaching task. The solid red curve is the mean of 10 runs with shaded area indicating the standard deviation. Comparison of the final energy cost with the ILQR-0 trajectory is plotted in the bar chart (**right**). The estimated input energy cost is 0.1843 ± 0.0204 J compared to ILQR-0's 0.2674 J.

The trajectories are parametrized by DMPs as introduced in Appendix D. Each sub-movement consists of 3 DMPs representing the trajectories of EP motor, stiffness motor and damping command. All DMP are initialized with shaping parameter $\mathbf{w} = \mathbf{0}$, which is a 10 dimensional vector. The goals $\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3$, for EP, stiffness motor, and damping respectively, are initialized as $\mathbf{g}_1 = \mathbf{q}^*$, $\mathbf{g}_2 = 24/\pi\ \mathbf{e}^{(3)}$, $\mathbf{g}_3 = 0.5\ \mathbf{e}^{(3)}$,[8] where $\mathbf{e}^{(d)}$ represents a d-dimensional unit

[8] $\mathbf{g}_2$ here is actually $\mathbf{p}_s$.

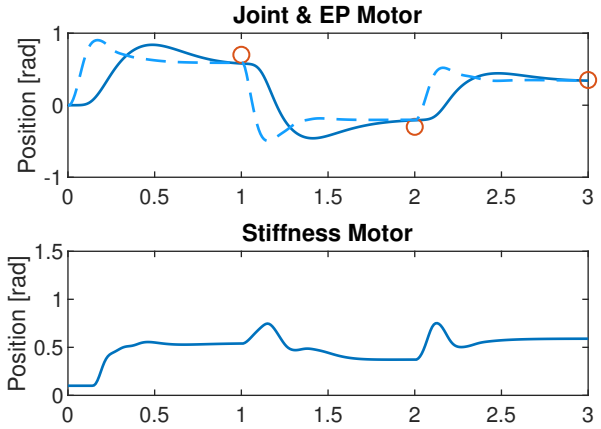


Figure 6.7: The final trajectory (one of 10 runs) of PI^2SEQ for the consecutive fast reaching task. (**Top**) Joint (solid line) and EP motor (dashed line) trajectories, while red dots denote the targets. (**Bottom**) Stiffness motor trajectory.

vector. The shaping parameter $\mathbf{w}$ is unconstrained. The box constraints are $[-\pi/3, \pi/3], [\pi/24, \pi/2], [0, 1]$ for elements of $\mathbf{g}_1$, $\mathbf{g}_2$ and $\mathbf{g}_3$ respectively. The overall policy parameter $\boldsymbol{\xi}$ is a 99-dimensional stacked vector of $\mathbf{g}$ and $\mathbf{w}$ of all three sub-movements. Relevant meta-parameters of the algorithm are $\gamma = 0.95, \mu = 15, K = 45$ and $\Sigma_{\boldsymbol{\epsilon}} = \text{diag}(10\ \mathbf{e}^{(90)}, 0.5\ \mathbf{e}^{(9)}) \in \mathbb{R}^{99\times 99}$, Both goals and shaping parameters $\mathbf{w}$ of DMPs are optimized simultaneously by Algorithm 1 except that it doesn't have an inner loop. The policy update rule used is the same as the weighted averaging method (6.10)-(6.12). Different from what suggested by Stulp et al. (2012), where the policy update takes the cost-to-go of sub-movements, we use the episodic cost along whole trajectory for policy update.[9]

The learning results of 10 sessions are illustrated in Figure 6.6. The policy update gradually converges and final energy cost evaluated by input mechanical

[9]By using cost-to-go for updating parameter of sub-movements, it presumes the reward/cost incurred during a sub-movement is independent from the sequent actions as argued in Stulp et al. (2012). However, another assumption underlying is that the *planning* of a action/parameter does not effect on the reward/cost prior to it. In fact, if this causality holds depends on if the planning happens at the time of action or at the very beginning. In classical RL an action is decided and executed at the same time, so the cost-to-go rather than total cost is used. While when RL is formulated as an optimization, the planning of all parameters takes place before all actions, the assumption underlying cost-to-go rule may not be true generally. Nevertheless, the choice is task-dependent and should be discussed for specific use-cases. In our practice, no obvious difference was found between two ways: (1) applying only cost-to-go to update parameters sequentially or (2) using total cost to update all parameters at once.

work E_{in} is successfully reduced from $0.2674\,\text{J}$ of ILQR-0 to $0.1843 \pm 0.0204\,\text{J}$. The final trajectory of one learning session shown in Figure 6.7 demonstrates that an optimized $\mathbf{p}_s$ regulates the pretension at the transition phases. The effectiveness of policy improvement with parametrized trajectories for exploiting variable impedance of VIAs is verified despite some drawbacks. First it can be seen that the learning takes thousands of (trajectory) samples due to high dimensionality of $\boldsymbol{\xi}$, which makes it less likely to be executed on the physical robot in an online fashion. Secondly, the joint trajectory in Figure 6.7 slightly but visibly deviated from the first two goals, because the integral term in (6.14) competes with the terminal cost of its previous movement. To circumvent this issue the composite cost function needs redesign to adjust the cost terms, weights, or impose extra constraints.

6.5.3 Task 1: sequential reaching with OC-ES

Now we take both weighting and stiffness parameters $\mathbf{w}_e, \mathbf{p}_s$ into account and employ the OC-ES framework. The policy parameter $\boldsymbol{\xi}$ for **Task 1** consists of weights of control effort term and stiffness motor preset of each sub-problem. $\boldsymbol{\xi}$ is initialized as $\{w_e^{(i)} = 1,\ p_s^{(i)} = \pi/24\,\text{rad}\}_{i=1}^{3}$. $K = 4$ roll-outs are run for each policy update up to 100 iterations. The exploration noise $\Sigma_{\boldsymbol{\epsilon}} = 0.5\ \mathbf{I} \in \mathbb{R}^{6\times 6}$ and decay factor γ is set to 0.95. The sample reuse parameter is chosen to be $\mu = 3$. The initial trajectory is evaluated to record its energy cost $E_{\text{in}}^{(0)}$ and $J_p^{(0)}$. The latter decides the upper bound constraint of reaching performance $\bar{J}_p$ with tolerance factor $\sigma_{\text{tol}} = 0.1$.

During each roll-out, $w_e^{(i)[k]}$ is used to set the weight of control effort term in (6.4) for i-th OCP, and $p_s^{(i)[k]}$ specifies the minimal position command $u_2^{\min}$ of the stiffness motor. By doing so, it constrains the minimal pretension upon reaching the target. The sub-problems are solved by ILQR. After each policy update, the movement without perturbation is evaluated to record the learning performance. To verify the improvement of energy saving, both initial and final trajectories are executed on the hardware to record the energy consumption. The results are summarized in Figure 6.8 where ILQR-ES denotes the final trajectory.

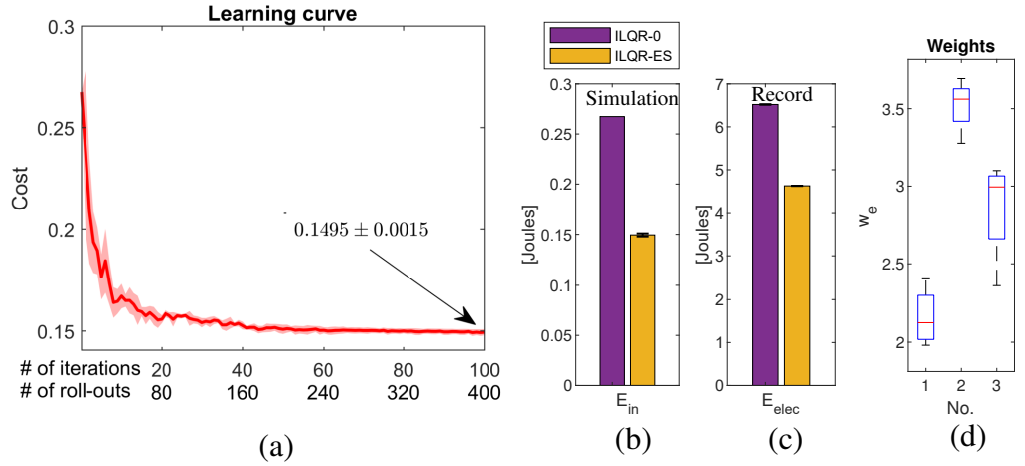


Figure 6.8: Shown are: (a) Learning curve of ILQR-ES for the consecutive fast reaching task. The solid red curve is the mean of 4 runs with shaded area indicating the standard deviation. (b) Estimated input energy cost of final result is 0.1495 ± 0.0015 J compared to 0.2674 J of ILQR-0. (c) Electrical consumption measured on servomotors is 6.5211 ± 0.2593 J (The gap between error bars is too small to be visible), while the benchmark ILQR-0 consumes 4.6261 ± 0.2812 J, which means a 29.6% reduction. (d) Distribution of optimal w_e for each sub–movement.

Significant energy reduction

The learning curve in Figure 6.8(a) shows a fast convergence after 50 iterations and very small variations as also shown in Figure 6.8(b) (c). It demonstrates that the OC-ES method successfully reduced the energy cost of the whole task whilst keeping worst performance cost within tolerance. The mechanical energy cost in simulation is decreased about 44% from that of the initial ILQR-0 trajectory. The electrical consumption recorded on the servomotors verifies the result with a 29.6% reduction.

Optimally tuned cost function weighting

Looking at Figure 6.8(d), all w_e of sub-problems tend to increase from the initial settings. Despite relative large variations of the optimization result, it can be observed that the second sub-movement takes the highest weight for control cost, which indicates the energy efficiency of it is most critical. In Figure 6.9 we can see

that the second sub-movement has the largest travel distance among the three, and consumes the most energy in the initial trajectory (as shown by the accumulated energy cost in Figure 6.9(i)). Hence, the result can be explained as the optimization adjusts the weight to balance the performance-cost trade-off more towards reducing energy cost.

Exploiting variable stiffness

It can be seen in Figure 6.9 that energy reduction occurs significantly during the second and third movements, compared with the initial trajectory. The stiffness motor maintains higher pretension at transition phases (Figure 6.9(f)) due to the constraint imposed by optimized $\mathbf{p}_s$, by which the acceleration of the subsequent movement consumes less energy in the EP motor. Also, the adjustment of stiffness motor causes a lot of electrical consumption (Figure 6.9(h)), suggesting that the control effort may be lead to suboptimal solutions regarding real energy consumption. However, this highly depends on the variable stiffness mechanism and hardware design. For example, by implementing the VSAs designed for minimizing energy cost for stiffness modulation (Jafari et al. (2015); Chalvet and Braun (2017)), the energy cost of the stiffness motor of initial trajectory can be reduced so that the most saving occurs on the EP motor. However, it would raise another problem that if a VSA does not require energy input to adjust stiffness, then it may not be able to pre-store energy at equilibrium position (EP). As a result there may be no energy buffering effect for some movements starting from a static equilibrium phase.

Overall, the experiment demonstrates the effectiveness of applying the OC-ES framework to improve energy efficiency by exploiting variable stiffness and cost function tuning. The learning takes only 4 explorations per iteration by leveraging model-based OC at the low-level, making it more feasible to run on the real robot.

6.5.4 Temporal and stiffness optimization for tracking control

The second application is to show that the proposed framework can be applied to temporal optimization and work with a low-level tracking controller.

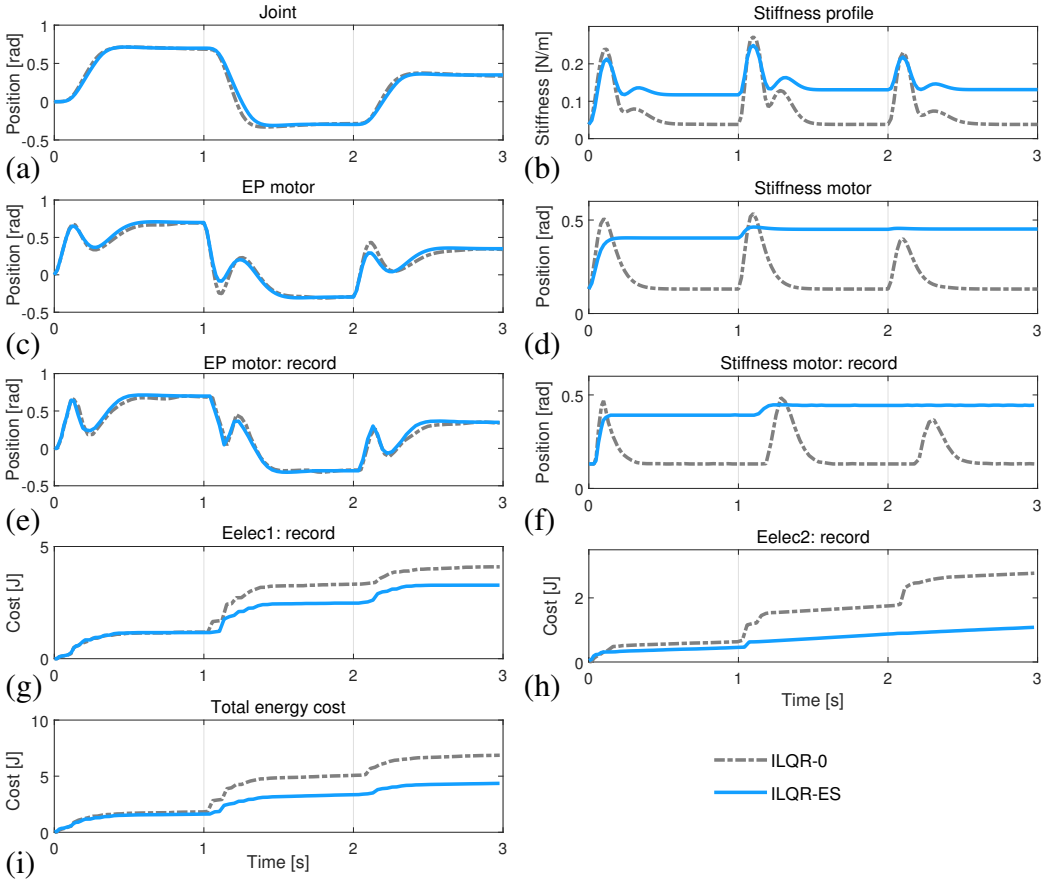


Figure 6.9: Result of executing the trajectory in simulation and on real hardware of ILQR-ES for consecutive fast reaching, compared with ILQR-0 trajectory as a benchmark. ILQR-0 also serves as the initial trajectory. The last two rows show the measured electrical cost by cumulating the recorded power along the trajectory.

Task 2 - Consecutive trajectory tracking

This task requires the arm to smoothly reach a sequence of targets with minimal-jerk joint trajectory. In addition to exploiting variable stiffness, the relative timing is allowed to be optimized. We set the targets as $\{q_i^*\} = [\pi/5, -0.2, 1, 0.3]$ (rad). The arm starts at $q(0) = 0\,\mathrm{rad}$. The total time for the movement is $2.4\,\mathrm{s}$. $\boldsymbol{\xi}$ is defined as $\boldsymbol{\xi} = (\mathbf{t}_d^\mathsf{T}, \mathbf{p}_s^\mathsf{T})^\mathsf{T} \in \mathbb{R}^7$, where $\mathbf{t}_d = \{t_d^{(i)}\}_{i=1}^3, \mathbf{p}_s = \{p_s^{(i)}\}_{i=1}^4$. Since the total time is kept the same, the last time duration is excluded from the policy

parameter. The box constraint on $\boldsymbol{\xi}$ is

$$\boldsymbol{\xi}_{\min} = (0.3, 0.3, 0.3, 0, 0, 0, 0)^{\mathsf{T}},\ \boldsymbol{\xi}_{\max} = (1.2, 1.2, 1.2, \frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2})^{\mathsf{T}}$$

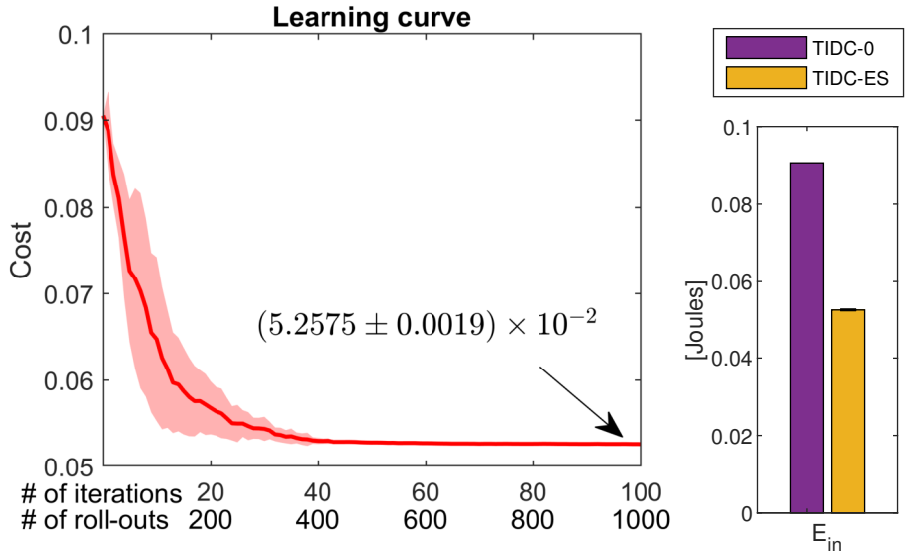


Figure 6.10: Learning curve (**left**) of TIDC-ES for the consecutive fast reaching task. The solid red curve is the mean of 10 runs with shaded area indicating the standard deviation. Comparison of the final energy cost with the TIDC-0 trajectory is plotted in the bar chart (**right**). The estimated input energy cost is $(5.2575 \pm 0.0019) \times 10^{-2}$ J compared to TIDC-0's 9.054×10^{-2} J.

The minimal-jerk joint trajectory can be computed analytically by formula introduced in Flash and Hogan (1985), given the time duration $\mathbf{t}_d$ and where it begins and ends. Then it becomes a joint space tracking problem. The joint tracking with extended inverse dynamics controller (TIDC) is derived in §3.5.1 to track the joint trajectory and resolve the actuation redundancy automatically. The controller serves as a feedback control law and reduces the inner loop OCP to a forward pass of dynamics. The stiffness parameter $\mathbf{p}_s$ is used to impose a constraint on the target position of stiffness motor in each sub-movement, by adding a null-space controller

$$\mathbf{v}_{\text{ns}} = ((q_i^* - \theta_1), p_s^{(i)} - \theta_2, 0)^{\mathsf{T}} \tag{6.16}$$

for i-th trajectory tracking. This null-space controller encourages the EP motor to move towards the joint target and the stiffness motor to $p_s^{(i)}$. Other relevant meta-parameters for the policy improvement method are: $K = 10, \mu = 3, \gamma = 0.97, \sigma_{\text{tol}} = 0.01, \Sigma_{\boldsymbol{\epsilon}} = \text{diag}(0.3\,\mathbf{e}^{(3)}, 0.5\,\mathbf{e}^{(4)})$. The whole method is termed as TIDC-ES.

Optimized parameters				
No.	1	2	3	4
$t_d^{(i)}$ [ms]	555.8 ± 5.1	593.1 ± 2.5	692.6 ± 3.4	558.5 ± 1.6
$p_s^{(i)}$[rad]	1.259 ± 0.013	0.782 ± 0.007	0.412 ± 0.008	0.002 ± 0.002

Table 6.1: Optimized parameters of temporal and stiffness optimization with TIDC-ES.

The initial trajectory is generated with $\boldsymbol{\xi}^{(0)} = (0.6, 0.6, 0.6, 0.2, 0.2, 0.2, 0.2)^\mathsf{T}$ and denoted by TIDC-0. The learning results after 100 iterations are presented in Figure 6.10. It can be seen that the learning curve initially has a large variation but quickly converges after 40 iterations. Compared to the initial trajectory, by exploiting stiffness and temporal optimization the input mechanical energy E_{in} reduces by about 42%. Looking at the results in Table 6.1, the optimized stiffness targets range from $1.26\,\text{rad}$ for the first sub-movement to nearly $0\,\text{rad}$ for the last one. It results in the pretension increasing during the first two sub-movements then decreasing towards the end (as shown in Figure 6.11). Moreover, the duration of third sub-movement is optimized to $692.6\,\text{ms}$, which is $92.6\,\text{ms}$ more than the initial setting, while other three sub-movements have shorter durations. The result is coherent with the order in terms of movement distance.

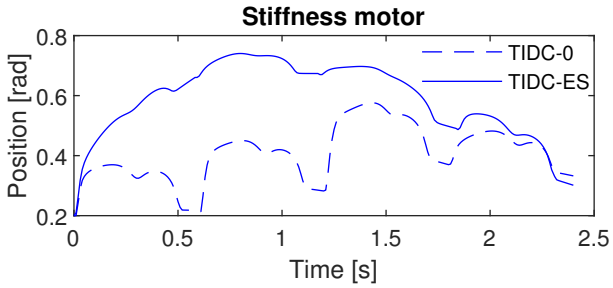


Figure 6.11: Stiffness motor profile of TIDC-ES final trajectory (blue solid) compared to initial trajectory (dashed blue).

6.5.5 Discussion

The experiments presented in this section demonstrated noticeable energy saving realized in consecutive reaching tasks. The task was the same as the one used in Chapter 5, although there a default spring pretension was chosen manually for all movements.

The proposed method in this chapter has been demonstrated to help the robot automatically regulate its stiffness with awareness of subsequent movements. The result of ILQR-ES for Task 1 regulates the stiffness motor to maintain at a small range around $0.5\,\mathrm{rad}$, suggesting that a fixed value can be tuned for energy efficiency in practice if the movement distances are not distributed diversely. In general, it suggests that for VIAs that rely on spring pretension to modulate stiffness, the more efficient way to use them in consecutive point-to-point reaching is not to reset their stiffness to minimum by default. Hence, using control effort to represent energy cost is questionable as it encourages the stiffness to move towards the minimum at the end of each submovement. However, due to the fact that accurate estimation of $E_{\mathrm{in}}, E_{\mathrm{elec}}$ needs a much smaller time step for discretization of the continuous system dynamics, the quadratic control effort is preferred for less computation cost. It also enhances smoothness of the trajectory, although at a cost to energetic optimality. Nevertheless, this loss is alleviated — by applying the proposed framework — with an upper layer optimizer that adjusts the trade-off balance.

6.6 Conclusions

This chapter proposed a versatile framework that integrates Optimal Control and Evolution Strategy (OC-ES) in a bi-level structure to address the optimization of movement sequence specifically for VIAs. At the low-level OC is leveraged to resolve the actuation redundancy and exploit variable impedance naturally. The high-level optimization in a sequential context is formulated as a reinforcement learning problem, in the form of iterative policy improvement, and solved as a black box optimization using method inspired by evolutionary strategy.

The proposed framework was applied for two consecutive reaching tasks on a

MACCEPA-VD actuator, one requires reaching as quickly as possible, the other tracks a smooth trajectory in joint space. In both cases natural dynamics is hard to be exploited for energy buffering as in periodic movement. By investigating the performance-cost trade-off via efficient frontiers, it can be seen how cost function weighting and minimal stiffness preset influence the energy efficiency. These two aspects can be addressed in the sequential movement context via the proposed framework, by which variable impedance can be fully exploited and the low-level trade-off is optimally balanced. In addition, a tracking controller that resolves the actuation redundancy was implemented to show the temporal and stiffness optimization at a high-level can improve the energy efficiency of low-level sequential tracking control. All the experiments presented in §6.5 demonstrated significant improvement of energy efficiency in both simulations and on hardware.

However, this work has been limited to reaching movements. More movement and task types need to be considered in the future work to demonstrate more complex behaviours. Furthermore, it is interesting to extend the application to compliant robots with multiple DOFs and consider problems involving contacts and interactions.

Chapter 7

Conclusions

7.1 Conclusions

This thesis investigated the design of a novel regenerative damping method to improve the energy efficiency of VIAs and established a framework for robots driven by variable impedance actuators (VIAs) to generate energy efficient sequential movements.

In Chapter 2 we reviewed the previous works that focus on optimizing the power flow of the elastic elements, however, the other major element in the whole power flow relation, *i.e.,* the dissipation via damping, has attracted little attention. We argued that this dissipated energy need not be lost, instead it can be harvested for later use. The approach proposed in this thesis, therefore, is to combine the roles of energy harvesting and damping, inspired by the regenerative braking technology used in vehicles. Also, it can be seen that most works in the VIA field have been constrained to rhythmic and discrete movements, there lacks a general method demonstrated for exploiting variable stiffness for sequential movements.

In Chapter 3 the optimal control framework is introduced. It started with the Euler-Lagrange derivation of system dynamics and concerns about simplification of motor dynamics. With an appropriate cost objective the optimal control problem can be solved by an efficient algorithm derived from differential dynamic programming. The actuation redundancy inherently involved in the VIAs is naturally resolved by optimal control. In addition, we provided another method to

cope with the redundancy, which is inversely mapping joint space targets to controls by using pseudo inverse. The resulting controller is able to track a given joint trajectory.

In Chapter 4 we proposed a novel regenerative damping design which consists of a DC motor attached to the robot joint and a dedicated control circuit to adjust the damping effect as required. The proposed damping module can be used to regenerate kinetic energy from bi-directional movements into a uni-directional power source, without reducing the maximum available damping. The control input for this damping module simply varies from 0 to 1, representing a proportional percentage of the maximum damping. As the power regeneration has a non-monotonic relation with the control input and damping coefficient (as verified by experiment), the balancing between damping allocation and energy regeneration needs to be treated with care. However, application of the hybrid damping module to VIAs in simulation, shows that it offers more flexibility to balance the trade-off between task performance and energy cost.

In Chapter 5, in order to investigate the use of variable regenerative damping for long-term operation, a stochastic consecutive reaching task was designed to examine the movement performance and energy efficiency. We designed metrics to evaluate the energy efficiency based on large sample dataset. Experiments are reported in which the VIA, equipped with the proposed damping module, performs sequential reaching to a series of stochastic targets. The results indicate that the combination of variable stiffness and variable regenerative damping results in a 25% performance improvement on metrics incorporating reaching accuracy, settling time, energy consumption and regeneration over comparable schemes where either stiffness or damping are fixed. The experimental study suggests that exploiting variable stiffness and variable damping simultaneously is desired, in such a way that the regenerative damping harvests kinetic energy when the simple "store and release" strategy of a spring can be exploited as in periodic movements. Also, there is more flexibility to prevent over-exploitation of variable damping and loss of regeneration capability.

A limitation of the experiment is that the "default" stiffness at the resting phase is determined by our experience to be high enough for stability of reaching on the real robot. However, from the point of view of sequential movement, the stiffness

at transition phases need not be fixed and should be optimally exploited.

In Chapter 6, we explored the problem of exploiting variable impedance for sequential movement. We proposed a hybrid approach that integrates optimal control with model-free reinforcement learning to enable learning of variable impedance skill for a sequence of movements. The framework leverages optimal control for resolving actuation redundancy and energy efficiency trade-off balance. The optimal control problems are then encapsulated in a reinforcement learning problem aiming at energy minimizing. The consecutive fast reaching task is then reused by allowing the optimization to fully exploit the variable impedance and control cost weighting. A second task was demonstrated to show its capability of dealing with optimization of temporal features, where the low-level smooth reaching problem utilizes the tracking controller we derived in Chapter 3 to track trajectories resulting from minimal jerk criterion. The experiments proved the effectiveness of the framework and successfully showed significant energy saving can be achieved by it in simulation and on hardware.

7.2 Recommendations for Future work

7.2.1 Electronics improvement

First, further prototyping of the power electronics integrated with the power management chip is required. The damping effect and control relationship needs to be identified in the presence of the power management chip and a real battery. Non-parametric model learning would be useful for fitting a model of regeneration estimation. Such information could be useful for high-level planning and control.

Also, the drawback of our current design is that the reflected inertia of the damping motor is increased by the gearbox and this adds to the link inertia. How to reduce this effect is an important question. The DC motor was used in our development to verify the principle, however, it is possible to develop the whole damping module without a DC motor. By doing this, the design parameter such as the torque constant, internal resistance, *etc.*, could be optimized and the inertia could be minimized.

7.2.2 Independent Damping controller

The regenerative damping scheme proposed in this thesis is modular and independent of specific VSA mechanism. Therefore, it can be integrated into any elastic joints and VSAs. The usefulness would be further justified by implementing on other VSA mechanisms.

A question is whether it is possible to design a modular damping controller as well so that the regenerative damping module is an off-of-shelf solution for users of VSAs. It would require investigation into the optimal control formulation and/or closed form controller to separate the control of damping from equilibrium position and stiffness. The control objective is relatively easy to define by *e.g.,* movement smoothness and stability. Due to the fact that damping control (using our regenerative damping) does not consume energy, it makes sense to take it out of the control effort and associate it with a cost function describing movement smoothness. However, whether a closed form controller is available or whether an optimization-based method should be used is still unknown. In the latter case the optimization should be easy and efficient to solve, ideally by the use of dynamic programming and analytical solution of the HJB equation.

7.2.3 Energy efficiency in the real world

In this thesis we discussed several aspects of energy efficiency of VIAs and conducted experiments with our 1-DOF MACCEPA. It would be important to generalize the discussion to higher DOF robots driven by VIAs. On those platforms it would be possible to identify movement tasks in the real world and compare with human movement data and other stiff robots. In such a way the use of energy regeneration can be evaluated and quantified to cross compare with the energy efficiency of stiff robots and even human beings.

The closed form inverse dynamics controller has been proven to be useful in the single DOF. Verification on high DOF robots needs to be done. For system with higher complexity, a closed form controller would be particularly useful and important for computation efficiency.

7.2.4 Parameterized movement representation

In the example of applying trajectory parametrization with DMPs, we implemented the vanilla DMP formulation which does not extend the generalization property of a DMP to VIA. This is because, although the shape of a single DMP is invariant in space and time, when movement duration and or goals (of EP, stiffness) shift, the resulting joint trajectory is not invariant. For example a rapid reaching would display a "lead and lag" in the EP motor domain, which means that the EP motor must lead to accelerate and then lag to decelerate. This phenomenon is commonly observed in VIAs and analogue to human arm reaching behaviour. The loss of invariant property would mean that the joint trajectory fails to adapt to suitable timing for "lead and lag". Moreover, another practical issue is that the dimensionality of the policy parameter is much higher than traditional stiff robots.

For this reason, parametric representation of VIA movement primitives should be more compact, invariant in joint space, and should encode the physical constraints.

Appendix A

Linear Algebra of Pseudo Inverse

When $\mathbf{A}$ is a square matrix of full rank, $\mathbf{Au} = \mathbf{b}$ has unique solution $\mathbf{u} = \mathbf{A}^{-1}\mathbf{b}$. However, if $\mathbf{A}$ is a m by n matrix where $m > n$, $\mathbf{A}$ is not invertible, the solution is solved by pseudo inverse. Moore-Penrose pseudo inverse is defined as $\mathbf{A}^\dagger = (\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T}$.

Lemma A.1. *The general solution of* $\mathbf{Au} = \mathbf{b}$ *when* $\mathbf{A} \in \mathbb{R}^{m\times n}, m > n$ *is*

$$\mathbf{u} = \mathbf{A}^\dagger\mathbf{b} + (\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{u}_1, \tag{A.1}$$

where $\mathbf{u}_1$ *is an arbitrary vector in the same space of* $\mathbf{u}$.

Note that if $rank(\mathbf{A}) = n$, $\mathbf{A}^\dagger = (\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T}$ and $\mathbf{A}^\dagger\mathbf{A} = \mathbf{I}$, so the second term vanishes and the solution becomes unique. In this case, $\mathbf{A}^\dagger$ is also unique.

Lemma A.2. *The solution* $\mathbf{u} = \mathbf{A}^\dagger\mathbf{b} + (\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{u}_1$ *that minimize* $J = \mathbf{u}^\mathsf{T}\mathbf{u}$ *is* $\mathbf{u} = \mathbf{A}^\dagger\mathbf{b}$.

Proof. Suppose we have $\mathbf{u} = \mathbf{A}^\dagger\mathbf{b}$ and $\mathbf{v} = \mathbf{u} + (\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{u}_1$.

$$\begin{aligned}
\mathbf{v}^\mathsf{T}\mathbf{v} = & \quad [\mathbf{A}^\dagger\mathbf{b} + (\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{u}_1]^\mathsf{T}[\mathbf{A}^\dagger\mathbf{b} + (\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{u}_1] \\
= & \quad (\mathbf{A}^\dagger\mathbf{b})^\mathsf{T}\mathbf{A}^\dagger\mathbf{b} + \mathbf{b}^\mathsf{T}(\mathbf{A}^\dagger)^\mathsf{T}(\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{u}_1 + \\
& ((\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{u}_1)^\mathsf{T}\mathbf{A}^\dagger\mathbf{b} + \mathbf{u}_1^\mathsf{T}(\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})^\mathsf{T}(\mathbf{I} - \mathbf{A}^\dagger\mathbf{A})\mathbf{u}_1.
\end{aligned} \tag{A.2}$$

Note that $\mathbf{A}^\dagger\mathbf{A}$ is symmetric (property of MP-inverse) we can show that in the

second term of (A.2),

$$\begin{aligned}(\mathbf{A}^\dagger)^\mathsf{T}(\mathbf{I}-\mathbf{A}^\dagger\mathbf{A}) &= (\mathbf{A}^\dagger)^\mathsf{T} - (\mathbf{A}^\dagger)^\mathsf{T}\mathbf{A}^\dagger\mathbf{A} \\ &= (\mathbf{A}^\dagger)^\mathsf{T} - (\mathbf{A}^\dagger)^\mathsf{T}(\mathbf{A}^\dagger\mathbf{A})^\mathsf{T} \\ &= (\mathbf{A}^\dagger)^\mathsf{T} - (\mathbf{A}^\dagger\mathbf{A}\mathbf{A}^\dagger)^\mathsf{T} \\ &= (\mathbf{A}^\dagger)^\mathsf{T} - (\mathbf{A}^\dagger)^\mathsf{T} = 0.\end{aligned}$$

Therefore, the second term and the third term (transpose of second term) both equal zero. Since the forth term is non-negative so it follows that

$$\mathbf{v}^\mathsf{T}\mathbf{v} - \mathbf{u}^\mathsf{T}\mathbf{u} = \mathbf{u}_1^\mathsf{T}(\mathbf{I}-\mathbf{A}^\dagger\mathbf{A})^\mathsf{T}(\mathbf{I}-\mathbf{A}^\dagger\mathbf{A})\mathbf{u}_1 \geq 0$$

$\therefore \mathbf{v} \geq \mathbf{u}$. □

Remark A.1. *Lemma A.2 means that the solution given by the MP-inverse is 'smaller' than any other solution, i.e., it has the smallest norm.*

Remark A.2. *In case the solution can not make the constraint strictly hold, the solution minimize the norm* $||\mathbf{A}\mathbf{u}-\mathbf{b}||$.

Lemma A.3. *The problem of*

$$\min J = \mathbf{u}^\mathsf{T}\mathbf{N}\mathbf{u} \tag{A.3}$$

$$s.t. \quad \mathbf{A}\mathbf{u} = \mathbf{b} \tag{A.4}$$

where $\mathbf{N}$ *is positive semi-definite matrix, has solution:* $\mathbf{u} = \mathbf{N}^{-\frac{1}{2}}(\mathbf{A}\mathbf{N}^{-\frac{1}{2}})^\dagger\mathbf{b}$.

Proof. Define $\mathbf{z} = \mathbf{N}^{\frac{1}{2}}\mathbf{u}$, where $\mathbf{N}^{\frac{1}{2}}\mathbf{N}^{\frac{1}{2}} = \mathbf{N}$, $\mathbf{N}$ is symmetric positive definite matrix, then $\mathbf{u} = \mathbf{N}^{-\frac{1}{2}}\mathbf{z}$. By substituting $\mathbf{u}$ it, we have

$$J = \mathbf{z}^\mathsf{T}(\mathbf{N}^{-\frac{1}{2}})^\mathsf{T}\mathbf{N}\mathbf{N}^{-\frac{1}{2}}\mathbf{z} = \mathbf{z}^\mathsf{T}\mathbf{z} \tag{A.5}$$

$$\mathbf{A}\mathbf{u} = \mathbf{A}\mathbf{N}^{-\frac{1}{2}}\mathbf{z} = \mathbf{b}. \tag{A.6}$$

The problem then becomes

$$\min J = \mathbf{z}^{\mathsf{T}}\mathbf{z} \tag{A.7}$$

$$s.t. \quad \mathbf{A}\mathbf{N}^{-\frac{1}{2}}\mathbf{z} = \mathbf{b} \tag{A.8}$$

According to Lemma A.2 The solution of this problem is

$$\mathbf{z} = (\mathbf{A}\mathbf{N}^{-\frac{1}{2}})^{\dagger}\mathbf{b}$$

$$\Rightarrow \mathbf{u} = \mathbf{N}^{-\frac{1}{2}}\mathbf{z} = \mathbf{N}^{-\frac{1}{2}}(\mathbf{A}\mathbf{N}^{-\frac{1}{2}})^{\dagger}\mathbf{b}. \tag{A.9}$$

□

Lemma A.4. *For* $\mathbf{N}$ *as a positive semi-definite matrix, the general solution of* $\mathbf{A}\mathbf{u} = \mathbf{b}$ *can be written as:*

$$\mathbf{u} = \mathbf{N}^{-\frac{1}{2}}(\mathbf{A}\mathbf{N}^{-\frac{1}{2}})^{\dagger}\mathbf{b} + \mathbf{N}^{-\frac{1}{2}}(\mathbf{I} - (\mathbf{A}\mathbf{N}^{-\frac{1}{2}})^{\dagger}\mathbf{A}\mathbf{N}^{-\frac{1}{2}})\mathbf{N}^{\frac{1}{2}}\mathbf{u}_1. \tag{A.10}$$

Proof. By assuming $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{N}^{-\frac{1}{2}}$ and $\mathbf{z} = \mathbf{N}^{\frac{1}{2}}\mathbf{u}$, we have $\tilde{\mathbf{A}}\mathbf{z} = \mathbf{b}$. According to A.1, the general solution is

$$\mathbf{z} = \tilde{\mathbf{A}}^{\dagger}\mathbf{b} + (\mathbf{I} - \tilde{\mathbf{A}}^{\dagger}\tilde{\mathbf{A}})\mathbf{z}_1. \tag{A.11}$$

By substituting back with $\mathbf{u}$ and $\mathbf{z}_1 = \mathbf{N}^{\frac{1}{2}}\mathbf{u}_1$, and multiply by $\mathbf{N}^{-\frac{1}{2}}$ on both side, the proof obviously completes. □

Appendix B

Motor Dynamics and Power Electronics

Modelling of DC motors

This section presents the modelling of DC motors, providing a foundation for model-based estimation of energy consumption. Also, we will and harvesting of DC motors.

In order to model the energy consumption,

Model

Both start with the fundamental dynamic equations based on Newton's 2nd law and Kirkoff's law:

$$J_m\ddot{\theta} + b_f\dot{\theta} = \tau_m - \tau_l \tag{B.1}$$

$$L\frac{\mathrm{d}i}{\mathrm{d}t} + R_m i = V_s - n_g k\dot{\theta} \tag{B.2}$$

where J_m is the motor inertia, b_f the viscous friction, L inductance, R_m resistance; V_s is power source voltage, i the current, τ_m, τ_l are motor torque and load torque; $\dot{\theta}, \ddot{\theta}$ represent the velocity and acceleration at the output shaft of gearbox. The equations and parameters are all reflected at the output shaft of gearbox, whose

gear reduction is denoted by n_g. The system (B.1) and (B.2), combined by the relation $\tau_m = n_g k i$ has order of three.

Energy consumption of DC motors

Following the basic mechanical and electrical equations (B.1) and (B.2) above, substituting $\tau_m = n_g k i$ into the equations yields:

$$P_{\text{elec}} = i^2 R_m + \tau_m \dot{\theta} + \frac{L\tau_m}{n_g^2 k^2} \frac{\mathrm{d}\tau_m}{\mathrm{d}t} \approx P_{\text{diss}}^{\text{e}} + P_{\text{mech}} \tag{B.3}$$

$$= i^2 R_m + J_m \ddot{\theta} \cdot \dot{\theta} + b_f \dot{\theta}^2 + \tau_l \dot{\theta} \tag{B.4}$$

$$= P_{\text{diss}}^{\text{e}} + P_{\text{mech}}^{\text{k}} + P_{\text{diss}}^{\text{m}} + P_{\text{load}} \tag{B.5}$$

$$P_{\text{diss}}^{\text{e}} = i^2 R_m \tag{B.6}$$

$$P_{\text{mech}} = \tau_m \dot{\theta} \tag{B.7}$$

$$P_{\text{mech}}^{\text{k}} = J_m \ddot{\theta} \cdot \dot{\theta} \tag{B.8}$$

$$P_{\text{diss}}^{\text{m}} = b_f \dot{\theta}^2 \tag{B.9}$$

$$P_{\text{load}} = \tau_l \dot{\theta} \tag{B.10}$$

Note that, for the input mechanical power P_{in} defined in Chapter 3, it is equivalent to P_{load}.

The derivative term with inductance is neglected because the time constant L/R_m of DC motor is normally very small. In the above equations, P_{elec} is the power of electrical energy consumption across the motor. $P_{\text{diss}}^{\text{e}}$ and $P_{\text{diss}}^{\text{m}}$ are dissipation on resistance and damping friction. From the equations above it can be seen that, the mechanical power P_{mech} consists of three components. Apart from the friction dissipation, $P_{\text{mech}}^{\text{k}}$ is the power flow associated with kinetic energy carried on the motor's rotor and gearbox, P_{load} represents the power delivered to, or received from the load. $P_{\text{elec}}, P_{\text{mech}}, P_{\text{load}}$ are all possible to be used to calculate energy consumption. If P_{elec} represents the most plausible one, then P_{mech} and P_{load} are two simplified versions. The terms overlooked by P_{mech} and P_{load} make these estimations problematic when they actually consume significant energy.

To investigate the relation between electrical power and mechanical power, equation (B.3) can be written as $P_{\text{elec}} = R_m \tau_m^2 / n_g^2 k^2 + \tau_m \dot{\theta}$. Figure B.1 plots

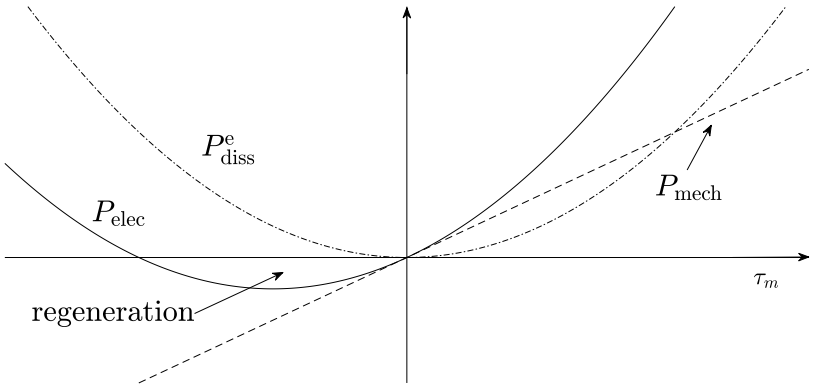


Figure B.1: Relation between P_{elec} and P_{mech}, w.r.t. the changes of motor torque τ_m, assuming velocity $\dot{\theta}$ is a positive value. When the motor delivers power to the load, $|P_{\mathrm{elec}}| \geqslant |P_{\mathrm{mech}}|$, while $|P_{\mathrm{elec}}| \leqslant |P_{\mathrm{mech}}|$ vice versa. The region of damping torque that the circuit can regenerate energy is $(-\dot{\theta} n_g^2 k^2/R, 0)$.

P_{elec}, P_{mech} and $P^{\mathrm{e}}_{\mathrm{diss}}$ as functions of τ_m (assuming $\dot{\theta}$ is a positive value), showing that during deceleration electrical energy may still be consumed although mechanical power is negative (negative torque represents braking against the velocity's direction). The region where energy is possible to be regenerated — by regenerative braking — is bounded by $(-\dot{\theta} n_g^2 k^2/R, 0)$. Noted that, since P_{elec} stands for the electrical power across the motor, when it has negative value it means that the motor works as a generator[1]. If running the motor as a generator is possible through the motor driver, the calculation of the electrical energy consumption can be made by taking integration of P_{elec} over time interval $[t_0, t_f]$:

$$E_{\mathrm{elec}} = \int_{t_0}^{t_f} P_{\mathrm{elec}} \mathrm{d}t. \tag{B.11}$$

While if regeneration is not allowed, then only the positive part of P_{elec} can be

[1]When motors work as generators, they have two modes: dynamical braking, regenerative braking and concurrent braking. The term dynamical braking is referred to the situation that the motor terminals are shorted, while regenerative braking is that the current sourced from motor charges to battery.

taken by E_{elec} (the operator $[\,]^+$ means simply treating all negative values as 0):

$$E_{\text{elec}} = \int_{t_0}^{t_f} [P_{\text{elec}}]^+ \mathrm{d}t. \tag{B.12}$$

In practice, calculation of (B.11) and (B.12) needs all the parameters involved well-identified. In case this condition is not available, the other energy consumptions may be used with carefulness in presence of energy regeneration on motor drivers; this will be discussed later in §C.

RC Servomotor

The analysis of dynamical model of DC servo motor shares the same method introduced above. On top of that, the controller of DC servo motor has to be taken into account for modelling the dynamical system and energy consumption. Modelling the energy consumption on controller would be straightforward if the user knows exactly the details of the controller. In this case, the DC motor system equations can be modified including the controller. We present a brief derivation as an example here, based on the assumption that the motor is controlled by a D-P controller, as identified by Wada et al. (2009) on the Hitec 7950TH servomotor. The formula of controlled input voltage reads:

$$V = K_p(\theta_d - \theta) - K_d\dot{\theta} \tag{B.13}$$

If we neglect electrical time constant[2], i.e., the transient response of current to input voltage is almost instantaneous, then tau_m is propositional to the input voltage V. The controller can be written as:

$$\tau_m = K_p(\theta_d - \theta) - K_d\dot{\theta} \tag{B.14}$$

Substituting the controller equation of τ_m into (B.1) leads to:

$$J_m\ddot{\theta} = K_p(\theta_d - \theta) - (K_d + b_f)\dot{\theta} - \tau_l \tag{B.15}$$

[2]preserving the derivative term of current will result in a system with one higher order.

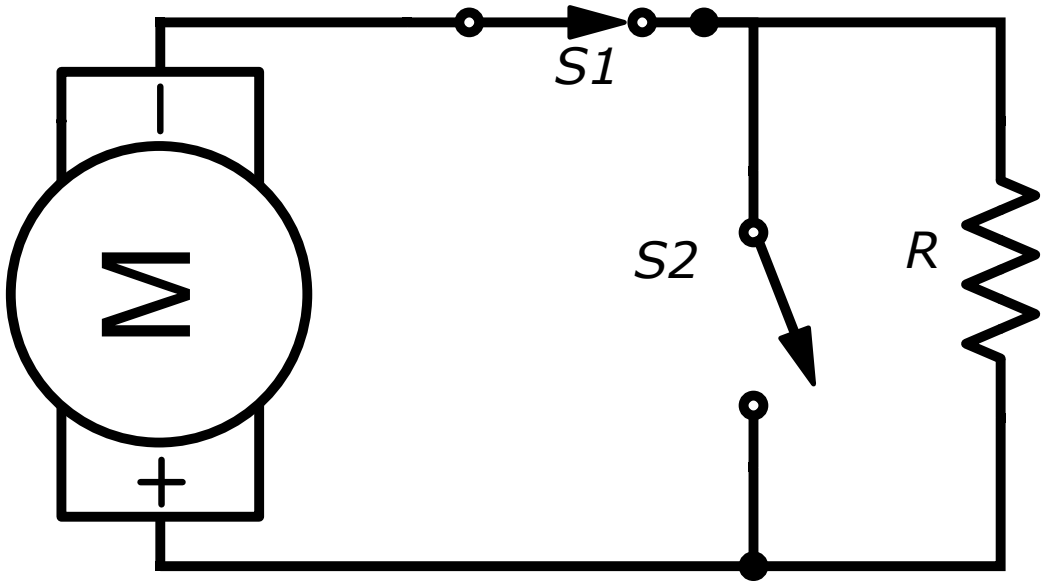


Figure B.2: Diagram representing basic operation of DC motor damping as a generator. S1 and S2 are two controlled switches, R represents an electric load that works as power sink and store electrical energy. When S1 switched on and S2 off, *regenerative braking* happens when motor M operates as generator and power flow into R; while when both S1, S2 switched on, no energy generated by M charges R and this is so-called *dynamical braking*. Another somehow dangerous case is *concurrent braking* when R also works as source and contributes to build up τ_m, resulting in large current and huge heat dissipation on the circuit. The switching frequency of S1 and S2 can be used to control the reflected damping property against mechanical load at motor's output shaft.

Energy regeneration through DC motor braking

As discussed above, electric DC motors can work as generators and recharge the energy from load to power source. This technique, often called regenerative braking, is common to be found on electrical cars. The MIT Cheetah, which can rival running animals in terms of efficiency, is a successful example of its application in robotics Seok et al. (2015). In contrast to implement regenerative braking directly on drive motors, this project examines the question of harvesting the dissipated energy on variable physical damping component. Figure B.2 shows a general diagram of the operation of DC motor regenerative braking and adjustable damping. Basically, there are two approaches as shown in the diagram, by controlling the

switches:

1. Scheme 1: remains S2 switched off, only regenerative braking is used. In this mode, S1 is switching for a certain duty-circle at high frequency. This duty-circle can be controlled to adjust the damping effect;
2. Scheme 2: both S1 and S2 are controlled to switch between dynamical braking (shorting the motor terminals) and regenerative braking. Dynamical braking generates higher damping, thus is desired when regenerative braking is not enough for required braking force.

Considering Scheme 1, assuming that a control input $u \in [0, 1]$ on S1 regulates the portion of time for which S1 is switched on; the electrical load in the circuit has resistance of R_l. The voltage generated across the motor reads $V_{\text{EMF}} = \eta(u) n_d k_b \dot{\theta}$, where n_d is the gear reduction of the motor, k_b is the speed constant, $\eta(u)$ modulates the voltage as a portion of the maximum voltage when S1 is always on[3]. The power flow charging R_l can be calculated by

$$P_r = \frac{R_l V_{\text{EMF}}}{R_m + R_l} \cdot \frac{V_{\text{EMF}}}{R_m + R_l} = \frac{R_l \eta(u)^2 n_d^2 k_b^2 \dot{\theta}^2}{(R_m + R_l)^2} \tag{B.16}$$

The damping torque is derived as

$$\tau_m = \frac{n_d k_t V_{\text{EMF}}}{R_m + R_l} = \frac{\eta(u) n_d^2 k_b^2 \dot{\theta}}{R_m + R_l} \tag{B.17}$$

and the damping coefficient reads

$$b_d(u) = \frac{\eta(u) n_d^2 k_b^2}{R_m + R_l} \tag{B.18}$$

It is obvious that the choice of R_l affects both the damping and power regeneration rate. By defining $\eta_r = R_l / R_m$ we can have

$$P_r = \frac{\eta_r}{(1 + \eta_r)^2} \frac{V_{\text{EMF}}^2}{R_m} \tag{B.19}$$

It can be seen that, η_r decides how much energy can be regenerated given a certain speed and control input u. $\eta_r / (1 + \eta_r)^2$ reaches its limit 0.5 at $\eta_r = 1$, which

[3]The simplest form of $\eta(u)$ is $\eta(u) = u$

simply suggests that for maximum energy regeneration R_l should be chosen to be equal to the motor resistance when designing the system. However, the bigger the load resistance R_l, the lower the damping effect. With $R_l = R_m$, the maximum damping achievable is halved from that of dynamical braking. This limitation can be overcome by implementing the Scheme 2, which switches between regenerative and dynamical braking, rather than free running, such that make up the loss of damping force. The development and implementation regarding this is discussed in Chapter 4. Here presents a simplified alternative:

$$b_d(u) = \frac{\eta(u) n_d^2 k_b^2}{R_m} \tag{B.20}$$

$$P_r = \alpha b_d(u) \dot{\theta}^2 \tag{B.21}$$

where α is the transmission efficiency from power generated through damping to the electrical load. It assumes that the limitation imposed by R_l is removed, and leaves $\eta(u)$ to be tuned.

Appendix C

The Role of Energetics in Optimal Control - A Numerical Study on MACCEPA

This section presents the preliminary study to investigate the role of energetic costs in optimal control and to predict the significance of energy regeneration. Particularly we are interested in the questions:

1. how the cost functions influence the shaping of optimal trajectory?
2. how much improvement in energy efficiency can be achieved by using the information about energy regeneration for optimal control?

To address the first question, we compare the optimization results of different cost criterion on a VIA actuated robot model — MACCEPA Van Ham et al. (2007) with variable damping (MACCEPA-VD) Radulescu et al. (2012). The model parameters are identified based on a hardware. Modelling of the energy regeneration and corresponding damping torque calculation are incorporated in this robot model. For the second question, numerical simulations were conducted to compare the trajectories optimized with or without regeneration term J_{rege}. To evaluate the effect of regenerative braking on energy optimality, some numerical experiments are reported here. The experimental methodology is as follows.

Minimal energetics

Minimal energy often refer to the cost criterion defined by the squared force and has a long history in optimal control of biological system Nelson (1983). Being a quadratic term, it also tends to eases the analytical analysis and prevails in robotics. The fact that the electrical dissipation is propositional to squared motor torque, as shown in Figure B.1, suggests that the squared motor torque is a good candidate when no regeneration on the motor driver. A list of the candidates of energy cost criterion is presented as follows,

$$\text{Control effort:} \qquad J_{\mathrm{u}} = \int_0^{t_f} (\mathbf{u} - \mathbf{u}^0)^{\mathsf{T}} \mathbf{H}_e (\mathbf{u} - \mathbf{u}^0)\, \mathrm{d}t \tag{C.1}$$

$$\text{Electrical energy:} \qquad J_{\mathrm{elec}} = \int_0^{t_f} \sum_{i=1}^{\mathrm{d_m}} [P^i_{\mathrm{elec}}]^+ \, \mathrm{d}t \tag{C.2}$$

$$\text{Mechanical energy:} \qquad J_{\mathrm{mech}} = \int_0^{t_f} \sum_{i=1}^{\mathrm{d_m}} [P^i_{\mathrm{mech}}]^+ \, \mathrm{d}t \tag{C.3}$$

$$\text{Mechanical output:} \qquad J_{\mathrm{load}} = \int_0^{t_f} \sum_{i=1}^{\mathrm{d_m}} [P^i_{\mathrm{load}}]^+ \, \mathrm{d}t \tag{C.4}$$

$$\text{Energy regeneration:} \qquad J_{\mathrm{rege}} = \int_0^{t_f} \sum_{i=1}^{\mathrm{n}} P_r \, \mathrm{d}t \tag{C.5}$$

An overall cost function can be defined to include the task cost J_t, energy cost, and regeneration:

$$J = J_t + w_e J_e - w_r J_{\mathrm{rege}} \tag{C.6}$$

The weighting parameter $w_e \geq 0$ can be adjusted to tune the trade-off between the primary task objective and energy consumption. J_e can be chosen from the cost criteria defined above. w_r can be modulated to reflect the importance of harvesting the energy; letting $w_r = w_e$ simply results that $J_e - J_{\mathrm{rege}}$, which is the estimation of net energy consumption.

Note that, J_{load} (C.4) may be of interest when compared with J_{rege} to reflect the energy regeneration efficiency because theoretically the mechanical work output from actuators is the source of energy regeneration via variable damping and defines the upper limit of the amount possibly to be harvested.

Control objective

The task considered in here is a point-to-point reaching movement. The task is encoded in the running cost as to minimize the position error integrated over time

$$J_t = \int_0^{t_f} \|\mathbf{q}(t) - \mathbf{q}^*\|^2 \, \mathrm{d}t \tag{C.7}$$

With this inserted in (C.6), the optimal control problem is to find optimal $\mathbf{u}$ that minimize the cost J defined by (C.6), subject to the robot dynamics model introduced above and initial state $\mathbf{x}(0) = \mathbf{x}_0$. The ILQR (Li and Todorov (2004)) method is implemented to solve this optimal control problem.

To address the two questions raised at the beginning of this section, we designed two numerical experiments.

Experiment 1 runs optimization using different cost criterion defined in §C and varies the weighting parameter w (assuming $w_r = w$) to investigate how cost criterion and their weighting affect the outcome of optimization. Comparison among these energy cost functions is made by plotting the optimal frontiers to illustrate the distribution of optimal solutions in cost-performance domain.

Experiment 2 uses J_{load} as the energy cost criterion, to compare the amount of energy regenerated from the plant. The more general form of energy regeneration model is used (B.20) (B.21) to put aside the limitation imposed by regenerative braking design. The parameters of the regeneration module, mainly n_d the motor gear reduction are tuned to make the maximum damping ability reach the level our hardware development targets at. Comparisons between optimal trajectories optimized with and without J_{rege} — denoted by $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ as comparing benchmark — are made to illustrate how the use of energy regeneration estimation can contribute to energy efficiency through optimal control. We ran the optimizations by varying both w and α, and iterated by 10 times.

The energy from actuators output onto the robot link, if not harvested or used for tasks including contact or manipulation, will be dissipated somewhere. The regeneration efficiency then can be defined in the sense of the amount of energy regenerated from the robot link: $\gamma = J_{\text{rege}}/J_{\text{load}}$. Since the task considered in this project is reaching movement, in this definition, we neglect the output work for

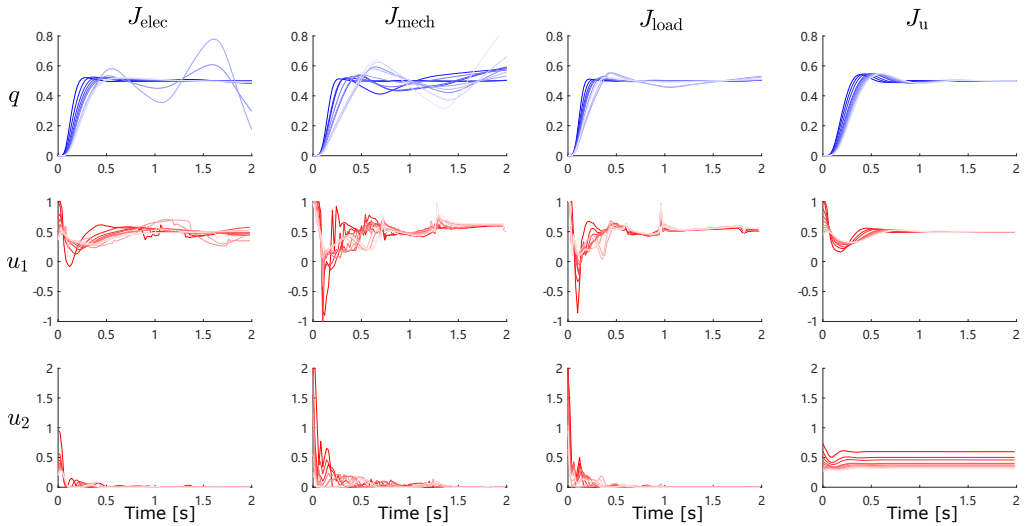


Figure C.1: Optimal trajectories of four energy cost criteria. Optimized with increasing w, shown by lightening colour. Top row of plots show the trajectory of joint angle q, middle and bottom rows illustrate the command sequences u_1 and u_2.

manipulating objects or contact with environment, and energy storage in springs. The reduction ratio,

$$\gamma_0 = 1 - \frac{J_{\text{load}}(\mathbf{x}, \mathbf{u}) - J_{\text{rege}}(\mathbf{x}, \mathbf{u})}{J_{\text{load}}(\hat{\mathbf{x}}, \hat{\mathbf{u}}) - J_{\text{rege}}(\hat{\mathbf{x}}, \hat{\mathbf{u}})} \tag{C.8}$$

is calculated by simulating each pair of trajectories for comparison and evaluating their energy loss of dissipation.

Results

In this section the results of numerical simulations are presented to show how the use of different energy cost influence the optimal solution. Furthermore, the role of energy regeneration in the context of optimal control is investigated.

Experiment 1: comparing different energy cost criterion

Optimizations and simulations were carried out as introduced in §C. In this experiment the robot was assumed to use only dynamical braking for variable damping adjustment. Figure C.1 illustrates the evolution of trajectories with increasing weighting w of the energy cost. The four groups share some similar features. For example, they tend to start with motor command for equilibrium point beyond the target position and extend the pretension simultaneously to elongate the spring, in such a way to generate large acceleration rate. When it needs braking, they reverse the equilibrium point to generate negative torque on the joint. With higher weighting for the energy cost, this strategy is less exploited because it generally costs a lot of energy on the motor side. Differences can be observed among these cost criterion that trajectories of J_{elec} and J_{u} exploit the above acceleration-deceleration strategy less than J_{mech} and J_{load}, and their motor commands are obviously smoother.

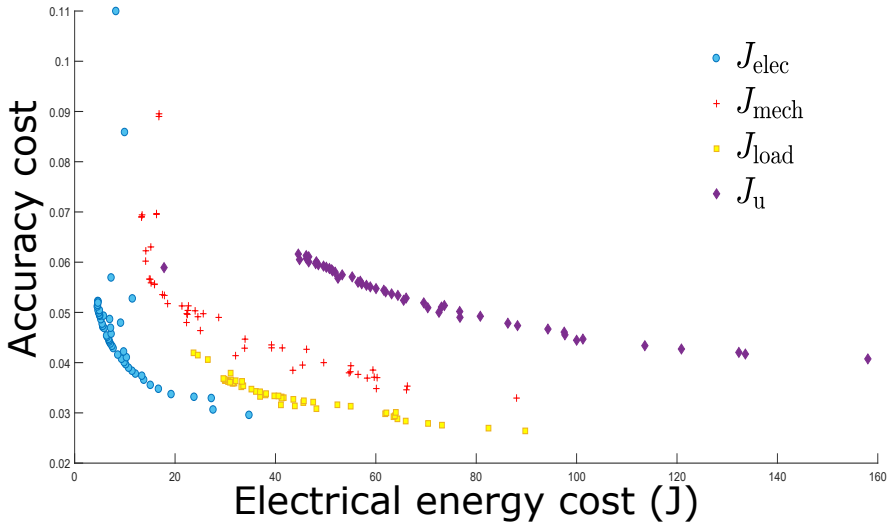


Figure C.2: Optimal frontiers of four energy cost criteria to show distribution of optimal trajectories in energy cost-task accuracy domain.

Although J_{u} results smoother commands and trajectories, it is defeated by others in the cost-performance domain. As shown in Figure C.2, the distribution of optimal frontier indicates that for the same energy cost, optimal trajectories of J_{u}

achieve the worst task performance, which is the accuracy cost encoded in the cost function for this reaching task experiment. The reason can be found in Figure C.1, as it shows that (i) the trajectories under J_{u} generally have consistent overshot while the other three groups can achieve smaller overshot with sufficiently small w; (ii) the pretension motor is heavily used and a lot of energy is stored by the spring. It is of interest to see if J_{u} is able to win over others in periodic tasks where energy storage in springs can be exploited to improve energy efficiency.

J_{elec} has the best performance in the optimal frontier plot with no surprise. However, it is interesting that J_{mech} does not show superior over J_{load} feature in this cost-performance domain, and the trajectories it results in have the largest sharp motor commands.

Experiment 2: the role of energy regeneration in optimal control

Energy regeneration of output mechanical work

The results of experiment 2 is presented here. Firstly, Figure C.3 — generated without taking energy regeneration estimation J_{rege} in optimisation — demonstrates a clear tendency that both the energy cost and its variation decreases significantly when using a relatively large weight parameter w. The output mechanical work is denoted in red colour, while the accuracy cost in blue. The box plots are used to show the variations (minimum, maximum, stand deviation) of 10 optimal solutions for a chosen value of weight w. In this set of simulations, as shown in Figure C.3 compared to $0.458\,J$ estimated output mechanical work, the cost reduced 61.6% to $0.176\,J$ at $w = 0.1$ quickly, and further reduction of 77.5% is achieved at $w = 0.3$. It demonstrates the energy regeneration design can achieve significant regeneration of energy carried on the plant. Admittedly, by increasing the gear reduction to amplify the damping range, it also increases friction between the joint and damping motor. Optimally balancing this trade-off according to user's preference needs a prudent choice of the cost function weightings and maybe an another layer of optimization.

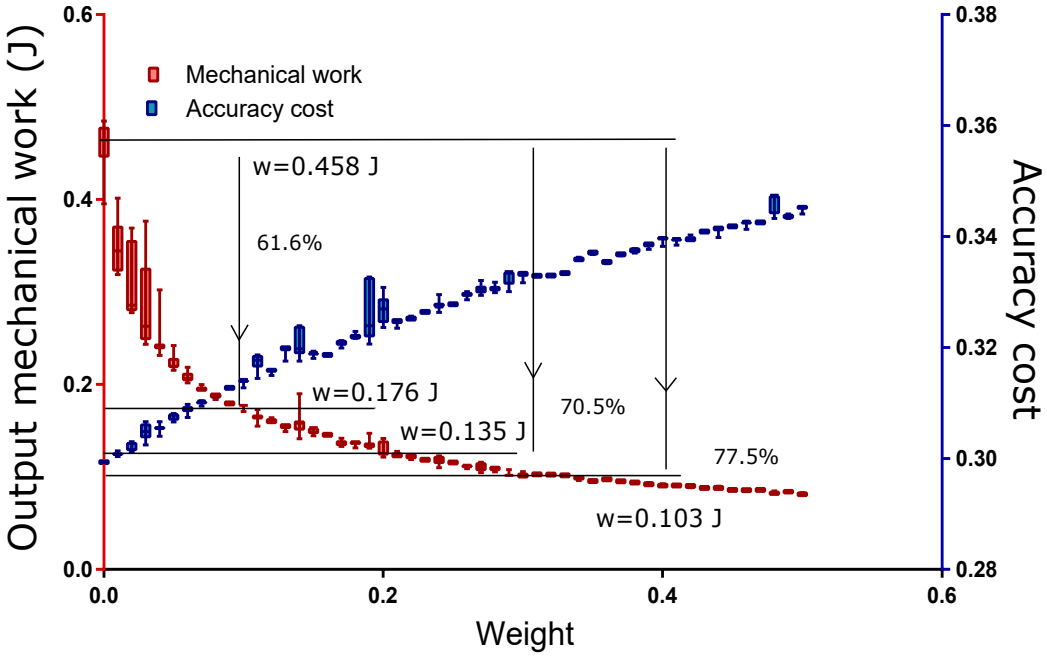


Figure C.3: Energy cost (red) and accuracy (blue) by varying weight parameter w from 0 to 0.5 with increment of 0.01. For each value of w, experiments were computed 10 times, whose maximum, minimum, and standard deviation are shown by box plot.

Incorporating energy harvesting information in optimisation

To further investigate how energy regeneration affects optimal trajectories when this information is taken into account, we compared the optimal trajectories optimized including J_{rege} with the benchmark trajectories $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$.

The reduction rate γ_0 defined by (C.8) are shown in Figure C.4. Noted that, since we ran the experiments for each combination of (w, α) 10 times, the data plotted in Figure C.4 are the means and standard deviations of γ_0. It can be seen that with a large regeneration transmission ratio α and for a certain chosen w, the optimization incorporating energy regeneration model results in a significant improvement for energy loss reduction. However, for smaller α — for example $\alpha = 0.5$, which is the limit of the regenerative braking scheme 1 — the reduction is not obvious. This may be because that, increasing damping in this case also increase dissipation in the damping circuit, thus the energy loss due to dissipation is not reduced.

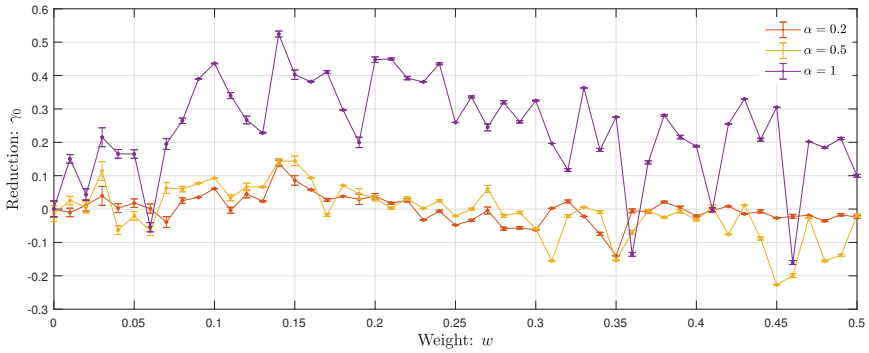


Figure C.4: Reduction rate γ_0, indicating the energy regeneration contributed by optimizing with J_{rege}. Trajectories are simulated on robot with different regeneration transmission ratio.

It has to be admitted that, since the weighting parameter w is tunable, there are cases where a trajectory optimized using a larger w without using J_{rege} costs even less energy than that have J_{rege} but with smaller w. Here we need optimal frontiers help to illustrate the overall difference caused optimization. From Figure C.5 we can see that when α is relatively small, there is no obvious energy reduction effect contributed by incorporating the information of energy regeneration in optimization, as we can hardly distinguish the two distributions. However, when α is relatively large (as $\alpha = 1$ in Figure C.5), the optimal frontier (in colour red) of trajectories shaped by incorporating the information of energy regeneration, is distributed below its counterpart (in colour blue) clearly, within a region roughly between $w = (0, 0.1)$. This represents that to achieve a certain performance level—in this work described by the accuracy cost—the former is more likely to consume less energy; and with a same energy constraint, it also tends to achieve higher performance.

In practice, when smoothness is important for safe trajectory execution, the traditional control effort is still useful and offers the possibility to balance performance-cost trade-off as well. Also, the when using the mechanical and electrical energetics, weight needs to be carefully chosen to avoid resulting undesired joint trajectory

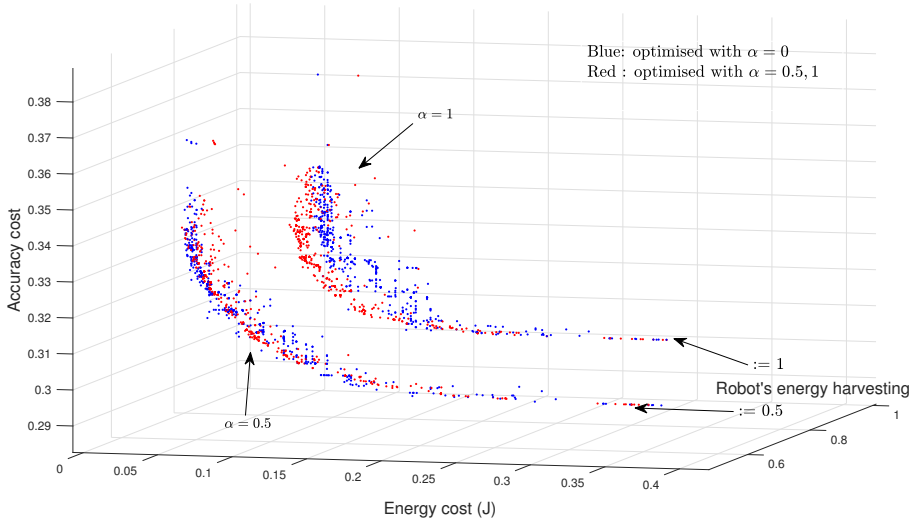


Figure C.5: Optimal frontiers show the distribution of optimal trajectories in cost-performance space. Blue points are the trajectories optimized without energy harvesting information (i.e, $\alpha = 0$ in cost function), and red points are the ones optimized w.r.t incorporating their energy harvesting ratios.

Appendix D

Trajectory Parametrization

The closed form tracking controller derived in §3.5.1 enables the possibility to generate a motion of VIAs by tracking a pre-stored trajectory, presumably parametrized to reduce the dimensionality of representation. A widely-used formalization is Dynamic Movement Primitive (DMP) proposed by Schaal (2006) and Ijspeert et al. (2002, 2013), based on the idea of dynamical system based modelling.

Dynamic Movement Primitives

Below is a formalization of DMPs for representing actuator variables $\boldsymbol{\theta}$.[1]

$$\tau\dot{\boldsymbol{\theta}} = \mathbf{z} \tag{D.1}$$

$$\tau\dot{\mathbf{z}} = \boldsymbol{\alpha}_z(\boldsymbol{\beta}_z(\mathbf{g} - \boldsymbol{\theta}) - \mathbf{z}) + s\mathbf{A}^{\mathsf{T}}\mathbf{f}^{\theta}(s) \tag{D.2}$$

$$\tau\dot{s} = -\alpha_s s \tag{D.3}$$

$$f_m^{\theta}(s) = \frac{\sum_{i=1}^{N}\psi_i(s)}{\sum_{i=1}^{N}\psi_i(s)} w_{m,i} \tag{D.4}$$

$$\psi_i(s) = \exp(-\frac{(s-c_i)^2}{2\sigma_i^2}) \tag{D.5}$$

[1] In our MACCEPA-VD case, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3)^{\mathsf{T}}$ represents the EP, stiffness motor and damping profile.

where $\tau > 0$ represents the duration and $\mathbf{g}$ is the goal position of $\boldsymbol{\theta}$, the dynamics of $\boldsymbol{\theta}$ is regulated by a dynamical system which behaves like a mass-spring-damper model, with gains determined by $\boldsymbol{\alpha}_z, \boldsymbol{\beta}_z$. $\mathbf{f}^{\theta}(s)$ is a forcing term manipulating the shape of the trajectory. It is a function in phase variable s, whose dynamics makes it asymptotically converge to 0, in a rate controlled by α_s. As a result, The efficacy of forcing term gradually decays to zero. This behaviour is purposely designed by Ijspeert et al. (2002, 2013) to enhance convergence of $\boldsymbol{\theta}$ to the goal $\mathbf{g}$. In addition to s, $\mathbf{A}$ is added to the forcing term to scale it according to the movement distance, where the m-th element $a_m = g_m - \theta_m$ corresponds to m-th forcing element f_m^{θ}. From (D.4) and (D.5) we can see that the forcing term is defined as the weighted sum of a set of N basis functions, of which each is an exponential function defined by centre point c_i and width factor σ_i.

DMP is a dynamical system with attractive point $\mathbf{g}$. The dynamical system can be transformed to a bounded-input, bounded-output (BIBO) system whose stability is easy to prove (Ijspeert et al. (2013)). Another appealing property of DMP is that it is invariant in space and time, in the sense that by scaling the distance and or time duration the "shape" of the path scales accordingly.

Trajectory optimization on basis of DMPs takes the shaping parameter $\mathbf{w}$ as policy parameters. The RL algorithm PI^2 (Theodorou et al. (2010b)) is commonly used to optimize it. As a DMP can be viewed as a controlled system linear in control vector $\mathbf{w}$, it is nicely suited to PI^2 — which is derived from stochastic optimal control. Stulp and Sigaud (2013) showed that, with some simplification, the algorithm coincides with the evolution strategy (ES). Our implementation is based on an evolutionary strategy and the insights given by Stulp *et al.* The policy improvement procedure is introduced in Chapter 6.

Appendix E

Hardware Development

Regenerative Damping Circuit

Our conceptual circuit was first investigated by simulation in Matlab Simulink to verify the effectiveness of the switching mechanism. Figure E.1 and Figure E.2 show the Simulink models of testing regenerative damping with a capacitor and a battery respectively. The power is generated by applying a velocity source at $5\,\mathrm{rad/s}$

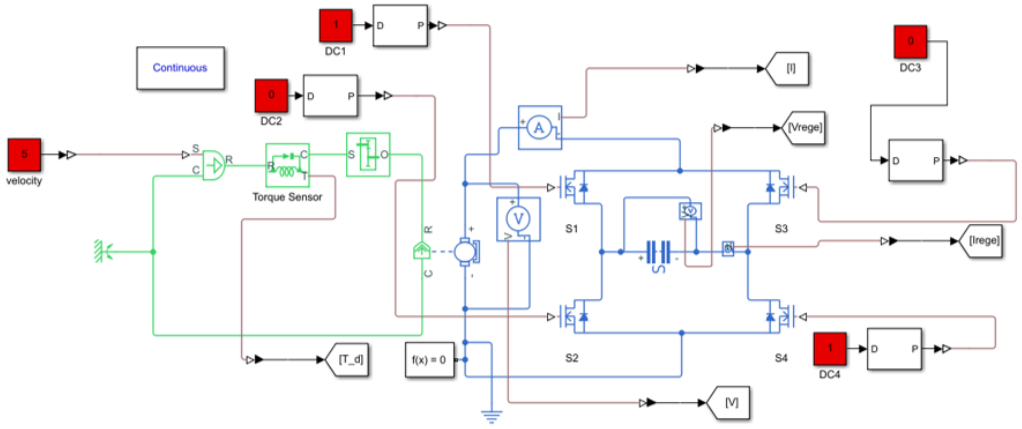


Figure E.1: Simulink model of testing regenerative damping with a capacitor.

With the help of Simulink modelling we verified the effectiveness of switching mechanism. Then a breadboard prototype was developed as shown in Figure E.3.

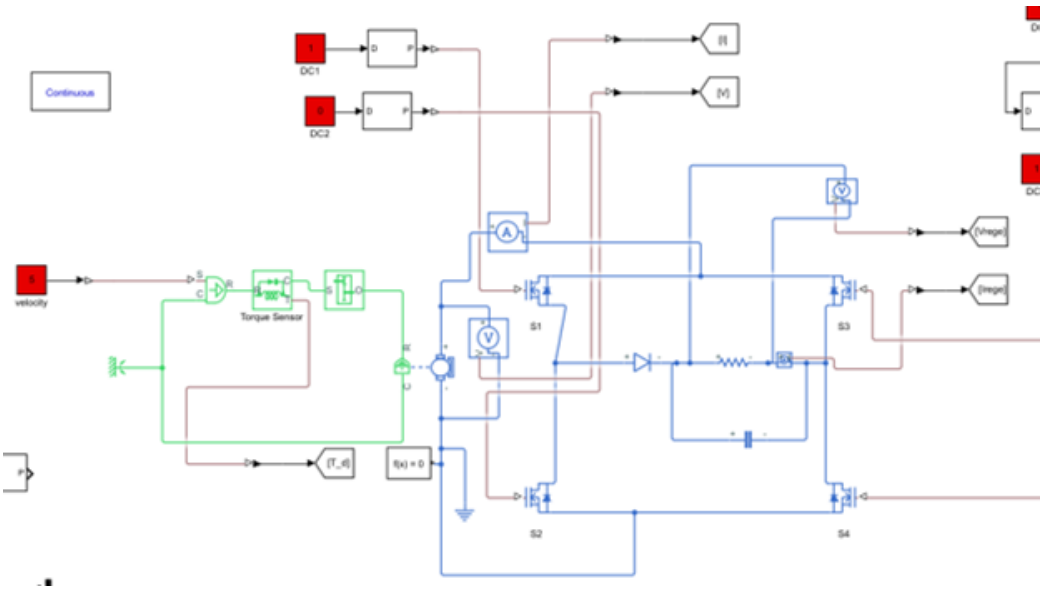


Figure E.2: Simulink model of testing regenerative damping with a battery.

This circuit prototype is used in the experiment presented in Chapter 4.

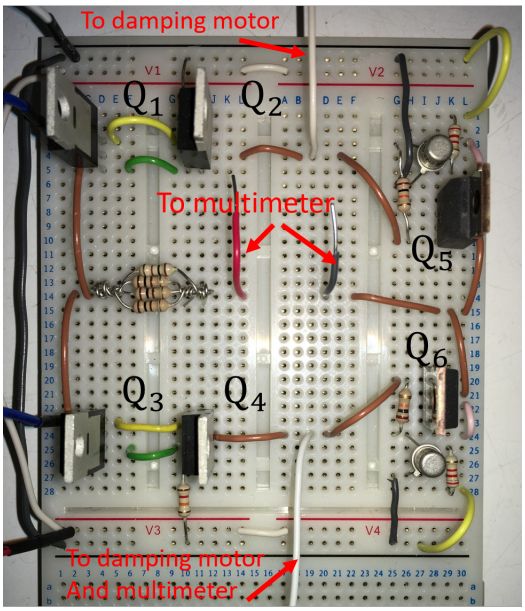


Figure E.3: Damping circuit prototype on a breadboard.

However, the regenerative damping circuit design do has drawback. Although the bi-directional function works passively, the PWM signals D_r, D_d are actively

controlled depending on the direction of current flow. The latency of current or velocity sensing would cause the control board fail to response immediately when movement direction reverses. This is not a critical issue for most smooth reaching movements. However, in case damping is required to suppress vibration, the sensing delay would make the realized damping effect deviate from the desired value. From practical view of point, in such vibration damping case the circuit can directly switch to dynamic braking mode.

Alternative circuit design with Uncontrolled bridge rectifier

The main reason MOSFETs are implemented for the regenerative damping control circuit is that the switch functionality of MOSFETs does not rely on triggering a threshold voltage. Commonly used uncontrolled bridge rectifier consists of four or more diodes. A diode typically has a threshold voltage that the voltage across it has to be higher than it to go through the diode. For example in our regenerative damping circuit integrated with power management chip, the diodes chosen for bridge rectifiers have a gating threshold of $1\,\mathrm{V}$. This means the back EMF voltage of damping motor has be greater than this threshold to be able to charge a power source. By experiments this circuit has been verified to be able to harvest energy from a constant voltage source. However, the characteristics of the damping effect has not been fully understood.

Sensors and actuators

Our 1-DOF MACCEPA was CAD designed and all parts except ball bearing were 3-D printed. The middleware is based on ROS system to send control commands and receive sensor records. In the development first version, the servomotors used are Hitec HS-7950TH for EP motor and HSR-5990TG for stiffness motor. A Maxon motor 110125 is chosen for damping motor. An Arduino Mega 2560 board is used to send PWM signals to the servos and reads sensor values from a joint potentiometer and current sensor Adafruit ina-219 measuring the power consumption of servos and regeneration power. The Arduino Mega board works as a ROS node and communicates with the PC. The PC runs the main ROS core

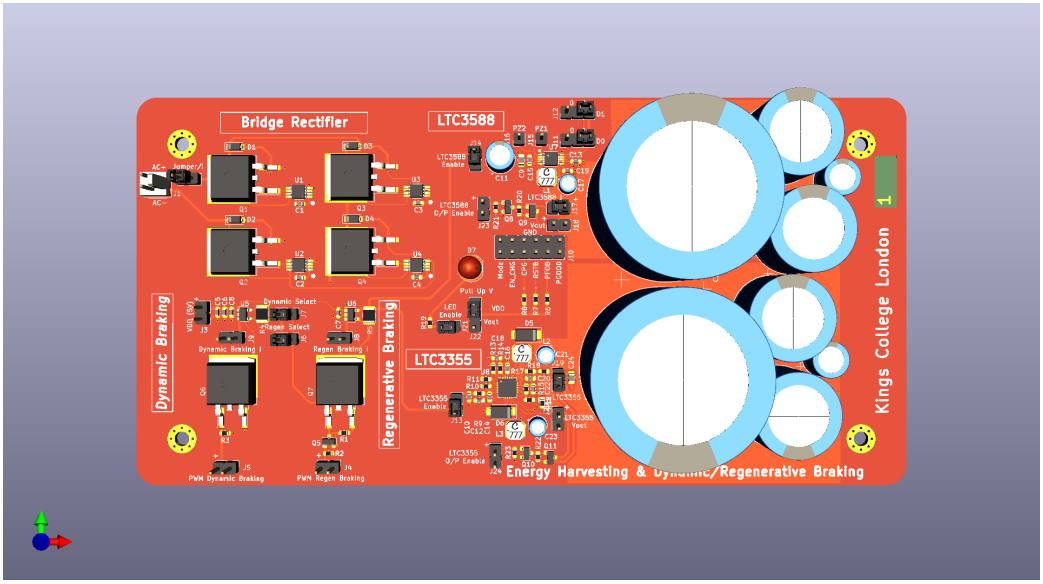


Figure E.4: PCB of regenerative damping circuit integrated with power management chip.

and several nodes that records sensor feedback and sends commands to Arduino. The experiments were all ran by Python scripts. Functions defined in Pythons are callable to interface with users using python script or Matlab program.

In the second development version, we replaced the PWM controlled servos with digital controlled servomotor Robotis Dynamixel XM430-W210,[1] to enable accurate sensing of servo current and position. The Dynamixel X series servos are capable of providing reliable position and current feedback. The data can be read transmitted using ROS system directly between the servos and PC with a dedicated U2D2 USB converter. We integrated the ROS SDK provided by Robotis to communicate between the servos and ROS based program.

[1] http://www.robotis.us/dynamixel-xm430-w210-r/

Bibliography

Albu-Schäffer, A., Eiberger, O., Fuchs, M., Grebenstein, M., Haddadin, S., Ott, C., Stemmer, A., Wimböck, T., Wolf, S., Borst, C. and Hirzinger, G. (2011). *Soft Robotics: From Torque Feedback Controlled Lightweight Robots to Intrinsically Compliant Systems*, volume 70 of *Springer Tracts in Advanced Robotics*. Berlin, Heidelberg, Springer Berlin Heidelberg.

Albu-Schäffer, A., Ott, C. and Hirzinger, G. (2007). A Unified Passivity-based Control Framework for Position, Torque and Impedance Control of Flexible Joint Robots. *The International Journal of Robotics Research* **26**(1): 23–39.

Alexander, R. M. (1988). *Elastic mechanisms in animal movement*. Wolters Kluwer Health.

Ariani, G. and Diedrichsen, J. (2019). Sequence learning is driven by improvements in motor planning. *Journal of Neurophysiology* **121**(6): 2088–2100.

Au, S. K., Weber, J. and Herr, H., Biomechanical Design of a Powered Ankle-Foot Prosthesis. In *IEEE 10th International Conference on Rehabilitation Robotics*. 2007, 298–303.

Bauml, B., Wimbock, T. and Hirzinger, G., Kinematically optimal catching a flying ball with a hand-arm-system. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2010, 2592–2599.

Baček, T., Moltedo, M., Geeroms, J., Vanderborght, B., Rodriguez-Guerrero, C. and Lefeber, D. (2020). Varying mechanical compliance benefits energy efficiency of a knee joint actuator. *Mechatronics* **66**: 102318.

Bellman, R. (1957). *Dynamic Programming*. Princeton University Press.

Berret, B., Ivaldi, S., Nori, F. and Sandini, G. (2011). Stochastic optimal control with variable impedance manipulators in presence of uncertainties and delayed feedback. *IEEE International Conference on Intelligent Robots and Systems* : 4354–4359.

Bobbert, M. F. (2001). Dependence of human squat jump performance on the series elastic compliance of the triceps surae: A simulation study. *Journal of Experimental Biology* **204**(3): 533–542.

Braun, D., Howard, M. and Vijayakumar, S. (2012). Optimal variable stiffness control: formulation and application to explosive movement tasks. *Autonomous Robots* **33**(3): 237–253.

Braun, D., Petit, F., Huber, F., Haddadin, S., Van Der Smagt, P., Albu-Schäffer, A. and Vijayakumar, S. (2013). Robots driven by compliant actuators: Optimal control under actuation constraints. *IEEE Transactions on Robotics* **29**(5): 1085–1101.

Braun, D. J., Apte, S., Adiyatov, O., Dahiya, A. and Hogan, N., Compliant actuation for energy efficient impedance modulation. In *IEEE International Conference on Robotics and Automation*. 2016, 636–641.

Braun, D. J., Chalvet, V. and Dahiya, A. (2018). Positive–Negative Stiffness Actuators. *IEEE Transactions on Robotics* **35**(1): 162–173.

Burdet, E., Osu, R., Franklin, D. W., Milner, T. E. and Kawato, M. (2001). The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature* **414**(6862): 446–449.

Burridge, R. R., Rizzi, A. A. and Koditschek, D. E. (1999). Sequential composition of dynamically dexterous robot behaviors. *International Journal of Robotics Research* **18**(6): 534–555.

Carabin, G., Wehrle, E. and Vidoni, R. (2017). A review on energy-saving optimization methods for robotic and automatic systems. *Robotics* **6**(4).

Catalano, M., Grioli, G., Garabini, M., Belo, F., Basco, A. d., Tsagarakis, N. and Bicchi, A., Implementing a variable impedance actuator. In *IEEE International Conference on Robotics and Automation*. IEEE, 2012.

Catalano, M. G., Grioli, G., Garabini, M., Bonomo, F., Mancini, M., Tsagarakis, N. and Bicchi, A., VSA-CubeBot: A modular variable stiffness platform for multiple degrees of freedom robots. In *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, 5090–5095.

Chalvet, V. and Braun, D. J. (2017). Criterion for the Design of Low-Power Variable Stiffness Mechanisms. *IEEE Transactions on Robotics* **33**(4): 1002–1010.

Cherelle, P., Matthys, A., Grosu, V., Vanderborght, B. and Lefeber, D., The AMP-Foot 2.0: Mimicking intact ankle behavior with a powered transtibial prosthesis. In *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*. IEEE, 2012, 544–549.

Collins, S. H. and Ruina, A., A bipedal walking robot with efficient and human-like gait. In *IEEE Int. Conf. Robotics & Automation*. 2005.

Collins, S. H., Ruina, A., Tedrake, R. and Wisse, M. (2005). Efficient Bipedal Robots Based on Passive-Dynamic Walkers. *Science* **307**(5712): 1082–1085.

Conn, A. R., Scheinberg, K. and Vicente, L. N. (2009). *Introduction to Derivative-Free Optimization*. SIAM.

Donelan, J. M., Li, Q., Naing, V., Hoffer, J. A., Weber, D. J. and Kuo, A. D. (2008). Biomechanical Energy Harvesting: Generating Electricity During Walking with Minimal User Effort. *Science* **319**(5864): 807–810.

Eiberger, O., Haddadin, S., Weis, M., Albu-Schaäffer, A. and Hirzinger, G., On joint design with intrinsic variable compliance: derivation of the DLR QA-Joint. In *2010 IEEE International Conference on Robotics and Automation*. IEEE, 2010, 1687–1694.

Faisal, A. A., Selen, L. P. J. and Wolpert, D. M. (2008). Noise in the nervous system. *Nature reviews. Neuroscience* **9**(4): 292–303.

Flash, T. and Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *The Journal of neuroscience* **5**(7): 1688–703.

Franklin, D. W., Burdet, E., Peng Tee, K., Osu, R., Chew, C.-M., Milner, T. E. and Kawato, M. (2008). CNS Learns Stable, Accurate, and Efficient Movements Using a Simple Algorithm. *Journal of Neuroscience* **28**(44): 11165–11173.

Furnemont, R., Mathijssen, G., van der Hoeven, T., Brackx, B., Lefeber, D. and Vanderborght, B., Torsion MACCEPA: A novel compact compliant actuator designed around the drive axis. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, 232–237.

Gao, Y., Chen, L. and Ehsani, M. (1999). Investigation of the Effectiveness of Regenerative Braking for EV and HEV. *SAE transactions* : 3184–3190.

Garabini, M., Passaglia, A., Belo, F., Salaris, P. and Bicchi, A., Optimality Principles in Variable Stiffness Control: The VSA Hammer. In *IROS*. 2011.

Goris, K., Saldien, J., Vanderborght, B. and Lefeber, D. (2011). MECHANICAL DESIGN OF THE HUGGABLE ROBOT PROBO. *International Journal of Humanoid Robotics* **08**(03): 481–511.

Grebenstein, M., Albu-Schaffer, A., Bahls, T., Chalon, M., Eiberger, O., Friedl, W., Gruber, R., Haddadin, S., Hagn, U., Haslinger, R., Hoppner, H., Jorg, S., Nickl, M., Nothhelfer, A., Petit, F., Reill, J., Seitz, N., Wimbock, T., Wolf, S., Wusthoff, T. and Hirzinger, G., The DLR hand arm system. In *IEEE Int. Conf. Robotics & Automation*. 2011.

Grimmer, M., Eslamy, M., Gliech, S. and Seyfarth, A., A comparison of parallel- and series elastic elements in an actuator for mimicking human ankle joint in walking and running. In *2012 IEEE International Conference on Robotics and Automation*. IEEE, 2012, 2463–2470.

Grimmer, M., Eslamy, M. and Seyfarth, A. (2014). Energetic and Peak Power Advantages of Series Elastic Actuators in an Actuated Prosthetic Leg for Walking and Running. *Actuators* **3**(1): 1–19.

Groothuis, S. S., Rusticelli, G., Zucchelli, A., Stramigioli, S. and Carloni, R. (2014). The variable stiffness actuator vsaUT-II: Mechanical design, modeling, and identification. *IEEE/ASME Transactions on Mechatronics* **19**(2): 589–597.

Guglielmino, E., Staimmers, C. W., Edge, K. A., Sireteanu, T. and Stancioiu, D. (2005). Damp-by-wire : Magnetorheological vs. friction damipers. *IFAC Proceedings Volumes (IFAC-PapersOnline)* **16**(July): 340–345.

Haddadin, S., Albu-Schäffer, A. and Hirzinger, G. (2009). Requirements for Safe Robots: Measurements, Analysis and New Insights. *The International Journal of Robotics Research* **28**(11-12): 1507–1527.

Haddadin, S., Huber, F. and Albu-Schaffer, A., Optimal control for exploiting the natural dynamics of Variable Stiffness robots. In *2012 IEEE International Conference on Robotics and Automation*. IEEE, 2012, 3347–3354.

Haddadin, S., Krieger, K., Albu-Schaffer, A. and Lilge, T. (2018). Exploiting Elastic Energy Storage for Blind Cyclic Manipulation: Modeling, Stability Analysis, Control, and Experiments for Dribbling. *IEEE Transactions on Robotics* **34**(1): 91–112.

Haddadin, S., Krieger, K., Kunze, M. and Albu-Schaffer, A., Exploiting potential energy storage for cyclic manipulation: An analysis for elastic dribbling with an anthropomorphic robot. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2011a.

Haddadin, S., Weis, M., Wolf, S. and Albu-schaeffer, A., Optimal control for maximizing link velocity of Robotic Variable Stiffness Joints. In *18th IFAC World Congress*. IEEE, 2011b.

Hansen, N. and Ostermeier, A. (2001). Completely Derandomized Self-Adaptation in Evolution Strategies. *Evolutionary Computation* **9**(2): 159–195.

Hellmund, R. E. (1917). Regenerative Braking of Electric Vehicles. *Transactions of the American Institute of Electrical Engineers* **XXXVI**: 1–78.

Hogan, N. (1985). Impedance control: An approach to manipulation: Part I-theory. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME* **107**(1): 17–24.

Hogan, N. and Buerger, S., Impedance and Interaction Control. In *Robotics and Automation Handbook*. CRC Press, 2004, 19–1.

Hogan, N. and Sternad, D. (2012). Dynamic primitives of motor behavior. *Biological Cybernetics* **106**(11-12): 727–739.

Howard, M., Braun, D. J. and Vijayakumar, S., Constraint-based equilibrium and stiffness control of variable stiffness actuators. In *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, 5554–5560.

Howard, M., Braun, D. J. and Vijayakumar, S. (2013). Transferring Human Impedance Behavior to Heterogeneous Variable Impedance Actuators. *IEEE Transactions on Robotics* **29**(4): 847–862.

Huang, H. J., Kram, R. and Ahmed, A. A. (2012). Reduction of Metabolic Cost during Motor Learning of Arm Reaching Dynamics. *The Journal of Neuroscience* **32**(6): 2182–2190.

Hutter, M., Remy, C. D., Hoepflinger, M. a. and Siegwart, R. (2013). Efficient and Versatile Locomotion With Highly Compliant Legs. *IEEE/ASME Transactions on Mechatronics* **18**(2): 449–458.

Ijspeert, A., Nakanishi, J. and Schaal, S. (2002). Movement imitation with nonlinear dynamical systems in humanoid robots. *IEEE International Conference on Robotics and Automation* (May): 1398–1403.

Ijspeert, A. J., Nakanishi, J., Hoffmann, H., Pastor, P. and Schaal, S. (2013). Dynamical movement primitives: learning attractor models for motor behaviors. *Neural computation* **25**(2): 328–73.

Ikeura, R. and Inooka, H., Variable Impedance Control fo a Robot for Cooperation with a Human. In *ICRA*. 1995, 3097–3102.

Jafari, A., Tsagarakis, N. and Caldwell, D. (2015). Energy efficient actuators with adjustable stiffness: a review on AwAS, AwAS-II and CompACT VSA changing stiffness based on lever mechanism. *Industrial Robot: An International Journal* **42**(3): 242–251.

Jafari, A., Tsagarakis, N. G. and Caldwell, D. G. (2013). Adjustable Stiffness (AwAS). *IEEE/ASME Transactions on Mechatronics* **18**(1): 355–365.

Jafari, A., Tsagarakis, N. G., Sardellitti, I. and Caldwell, D. G. (2014). A New Actuator With Adjustable Stiffness Based on a Variable Ratio Lever Mechanism. *IEEE/ASME Transactions on Mechatronics* **19**(1): 55–63.

Jimenez-Fabian, R., Geeroms, J., Flynn, L., Vanderborght, B. and Lefeber, D. (2017). Reduction of the torque requirements of an active ankle prosthesis using a parallel spring. *Robotics and Autonomous Systems* **92**: 187 – 196.

Jäntsch, M., Wittmeier, S., Dalamagkidis, K., Herrmann, G. and Knoll, A., Adaptive neural network Dynamic Surface Control: An evaluation on the musculoskeletal robot Anthrob. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*. 2015, 4347–4352.

Jäntsch, M., Wittmeier, S., Dalamagkidis, K., Panos, A., Volkart, F. and Knoll, A., Anthrob - A printed anthropomimetic robot. In *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*. 2013, 342–347.

Kashiri, N., Abate, A., Abram, S. J., Albu-Schaffer, A., Clary, P. J., Daley, M., Faraji, S., Furnemont, R., Garabini, M., Geyer, H., Grabowski, A. M., Hurst, J., Malzahn, J., Mathijssen, G., Remy, D., Roozing, W., Shahbazi, M., Simha, S. N., Song, J. B., Smit-Anseeuw, N., Stramigioli, S., Vanderborght, B., Yesilevskiy, Y. and Tsagarakis, N. (2018). An overview on principles for energy efficient robot locomotion. *Frontiers Robotics AI* **5**(DEC).

Khatib, O. (1987). A unified approach for motion and force control of robot manipulators: the operational space formulation. *IEEE Journal ofRobotics and Automation* **3**(1): 43–55.

Kirk, D. E. (1970). *Optimal Control Theory: An Introduction*. NewYork: Prentice-Hall.

Kong, K., Bae, J. and Tomizuka, M. (2009). Control of Rotary Series Elastic Actuator for Ideal Force-Mode Actuation in Human–Robot Interaction Applications. *IEEE/ASME Transactions on Mechatronics* **14**(1): 105–118.

Laffranchi, M., Chen, L., Tsagarakis, N. G. and Caldwell, D. G., The role of physical damping in compliant actuation systems. In *IEEE Int. Conf. Intelligent Robots & Systems*. 2012a.

Laffranchi, M., Tsagarakis, N. G. and Caldwell, D. G. (2010). A variable physical damping actuator (VPDA) for compliant robotic joints. *Proceedings - IEEE International Conference on Robotics and Automation* : 1668–1674.

Laffranchi, M., Tsagarakis, N. G. and Caldwell, D. G., CompAct™ Arm: a Compliant Manipulator with Intrinsic Variable Physical Damping. In *Robotics: Science and Systems*. 2012b.

Laffranchi, M., Tsagarakis, N. G. and Caldwell, D. G., Improving Safety of Human-Robot Interaction Through Energy Regulation Control and Passive Compliant Design. In *Human Machine Interaction-Getting Closer*. 2012c.

Laffranchi, M., Tsagarakis, N. G. and Caldwell, D. G. (2012d). Variable physical damping actuators (VPDAs): Facilitating the control and improving the performance of compliant actuation systems. *International Workshop on Advanced Motion Control, AMC* .

Laffranchi, M., Tsagarakis, N. G. and Caldwell, D. G. (2013). Analysis and Development of a Semiactive Damper for Compliant Actuation Systems. *IEEE/ASME Transactions on Mechatronics* **18**(2): 744–753.

Lakatos, D., Petit, F. and Albu-Schäffer, A. (2014). Nonlinear Oscillations for Cyclic Movements in Human and Robotic Arms. *IEEE Transactions on Robotics* **30**(4): 865–879.

Lashley, K. S., The problem of serial order in behavior. In *Cerebral mechanisms in behavior; the Hixon Symposium.* Oxford, England: Wiley, 1951, 112–146.

Lay, B., Sparrow, W., Hughes, K. and O'Dwyer, N. (2002). Practice effects on coordination and control , metabolic energy expenditure , and muscle activation. *Human Movement Science* **21**: 807–830.

Levine, S. and Koltun, V. (2012). Continuous inverse optimal control with locally optimal examples. *Proceedings of the 29th International Conference on Machine Learning, ICML 2012* **1**: 41–48. 1206.4617.

Li, W. and Todorov, E., Iterative Linear Quadratic Regulator Design for Nonlinear Biological Movement Systems. In *IEEE Int. Conf. Robotics & Automation*. 2004.

Li, W. and Todorov, E. (2007). Iterative linearization methods for approximately optimal control and estimation of non-linear stochastic system. *International Journal of Control* **80**(9): 1439–1453.

Lucia, P., Umezawa, K., Nakamura, Y. and Billard, A., Learning Robot Skills Through Motion Segmentation and Constraints Extraction. In *HRI Workshop on Collaborative Manipulation*. 2013.

Lynch, K. M. and Park, F. C. (2017). *Modern Robotics: Mechanics, Planning, and Control*. Cambridge University Press.

Manschitz, S., Kober, J., Gienger, M. and Peters, J. (2015). Learning movement primitive attractor goals and sequential skills from kinesthetic demonstrations. *Robotics and Autonomous Systems* **74**: 97–107.

Matsusaka, K., Uemura, M. and Kawamura, S. (2016). Realization of highly energy efficient pick-and-place tasks using resonance-based robot motion control. *Advanced Robotics* **30**(9): 608–620.

Maufroy, C., Siee, M., Hoepf, M., Verl, A., Lenz, A. and Dalamagkidis, K., MYOROBOTICS : A modular toolkit for musculoskeletal robot development. In *ICRA Workshop on Research Challenges and Cross-cutting Applications of Modular Robot Systems*. 2014.

McCallion, H., Johnson, G. and Pham, D. (1979). A compliant device for inserting a peg in a hole. *Industrial Robot: An International Journal* **6**(2): 81–87.

Milner, T. E. and Cloutier, C. (1998). Damping of the wrist joint during voluntary movement. *Experimental Brain Research* **122**(3): 309–317.

Mombaur, K., Truong, A. and Laumond, J.-P. (2010). From human to humanoid locomotion—an inverse optimal control approach. *Autonomous Robots* **28**(3): 369–383.

Morita, T. and Sugano, S., Design and development of a new robot joint using a mechanical impedance adjuster. In *Proceedings of 1995 IEEE International Conference on Robotics and Automation*, volume 3. IEEE, 1995, 2469–2475.

Motegi, M. and Matsui, T. (2011). Optimal Control Model for Reproducing Squat Movements Based on Successive-Movement Combination. *The Proceedings of the Symposium on sports and human dynamics* **2011**: 558–563.

Nakamura, Y., Hanafusa, H. and Yoshikawa, T. (1987). Task-Priority Based Redundancy Control of Robot Manipulators. *The International Journal of Robotics Research* .

Nakanishi, J., Radulescu, A., Braun, D. J. and Vijayakumar, S. (2016). Spatio-temporal stiffness optimization with switching dynamics. *Autonomous Robots* : 1–19.

Nakanishi, J., Rawlik, K. and Vijayakumar, S., Stiffness and temporal optimization in periodic movements: An optimal control approach. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1. 2011, 718–724.

Nelson, W. L. (1983). Physical principles for economies of skilled movements. *Biological cybernetics* **46**: 135–147.

Okada, M., Ban, S. and Nakamura, Y., Skill of compliance with controlled charging/discharging of kinetic energy. In *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, volume 3. IEEE, 2002, 2455–2460.

Ozparpucu, M. C. and Haddadin, S., Optimal control for maximizing link velocity of visco-elastic joints. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2013, 3035–3042.

Ozparpucu, M. C. and Haddadin, S., Optimal control of elastic joints with variable damping. In *2014 European Control Conference (ECC)*, 4. IEEE, 2014, 2526–2533.

Paluska, D. and Herr, H. (2006). The effect of series elasticity on actuator power and work output: Implications for robotic and prosthetic joint design. *Robotics and Autonomous Systems* **54**(8): 667–673.

Pandy, M. G., Zajac, F. E., Sim, E. and Levine, W. S. (1990). An optimal control model for maximum-height human jumping. *Journal of Biomechanics* **23**(12): 1185–1198.

Paryanto, Brossog, M., Bornschlegl, M. and Franke, J. (2015). Reducing the energy consumption of industrial robots in manufacturing systems. *International Journal of Advanced Manufacturing Technology* **78**(5-8): 1315–1328.

Pellicciari, M., Avotins, A., Bengtsson, K., Berselli, G., Bey, N., Lennartson, B. and Meike, D., AREUS - Innovative hardware and software for sustainable industrial robotics. In *IEEE International Conference on Automation Science and Engineering*, volume 2015-Octob. 2015, 1325–1332.

Pfeifer, R. and Bongard, J. (2007). *How the Body Shapes the Way We Think: A New View of Intelligence*. The MIT Press.

Pitou, S., Wu, F., Shafti, A., Michael, B., Stopforth, R. and Howard, M., Embroidered Electrodes for Control of Affordable Myoelectric Prostheses. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, 1812–1817.

Plooij, M., Wisse, M. and Vallery, H. (2016). Reducing the Energy Consumption of Robots Using the Bidirectional Clutched Parallel Elastic Actuator. *IEEE Transactions on Robotics* **32**(6): 1512–1523.

Pratt, G. A. and Williamson, M. M., Series elastic actuators. In *1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. 'Human Robot Interaction and Cooperative Robots'*, volume 1. 1995, 399–406.

Pratt, J. E. (2000). *Exploiting Inherent Robustness and Natural Dynamics in the Control of Bipedal Walking Robots*. PhD Thesis.

Radulescu, A. (2016). *Exploiting Variable Impedance in domains with contacts*. PhD Thesis, University of Edinburgh.

Radulescu, A., Howard, M., Braun, D. J. and Vijayakumar, S., Exploiting variable physical damping in rapid movement tasks. In *IEEE/ASME Int. Conf. Advanced Intelligent Mechatronics*. 2012.

Raibert, M. H. (1986). *Legged Robots That Balance*. Cambridge, MA, USA, Massachusetts Institute of Technology.

Rawlik, K., Toussaint, M. and Vijayakumar, S. (2010). An Approximate Inference Approach to Temporal Optimization in Optimal Control. *Neural Information Processing Systems* : 1–9.

Reher, J., Cousineau, E. A., Hereid, A., Hubicki, C. M. and Ames, A. D., Realizing dynamic and efficient bipedal locomotion on the humanoid robot DURUS. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*. 2016, 1794–1801.

Remy, C. D., Buffinton, K. and Siegwart, R., Comparison of cost functions for electrically driven running robots. In *2012 IEEE International Conference on Robotics and Automation*. IEEE, 2012, 2343–2350.

Richter, C., Jentzsch, S., Hostettler, R., Garrido, J. A., Ros, E., Knoll, A., Rohrbein, F., van der Smagt, P. and Conradt, J. (2016). Musculoskeletal Robots: Scalability in Neural Control. *IEEE Robotics & Automation Magazine* **23**(4): 128–137.

Roberts, T. J. (2016). Contribution of elastic tissues to the mechanics and energetics of muscle function during movement. *The Journal of Experimental Biology* **219**(2): 266–275.

Robinson, D. W., Pratt, J. E., Paluska, D. J. and Pratt, G. A., Series Elastic Actuator Development for a Biometric Walking Robot. In *Proceedings of the*

1999 IEEE/ASME International Conference on Advanced Intelligent Machines. 1999, 561–568.

Roozing, W., Li, Z., Caldwell, D. G. and Tsagarakis, N. G. (2016). Design Optimisation and Control of Compliant Actuation Arrangements in Articulated Robots for Improved Energy Efficiency. *IEEE Robotics and Automation Letters* **1**(2).

Roozing, W., Ren, Z. and Tsagarakis, N. G. (2019). An efficient leg with series–parallel and biarticular compliant actuation: design optimization, modeling, and control of the eLeg. *International Journal of Robotics Research* .

Sakagami, Y., Watanabe, R., Aoyama, C., Matsunaga, S., Higaki, N. and Fujimura, K., The intelligent ASIMO: system overview and integration. In *IEEE/RSJ International Conference on Intelligent Robots and System*, volume 3. IEEE, 2002, 2478–2483.

Salisbury, K., Eberman, B., Levin, M. and Townsend, W., The design and control of an experimental whole-arm manipulator. In *The fifth international symposium on Robotics research*. 1991, 233–241.

Schaal, S. (2006). Dynamic Movement Primitives – A Framework for Motor Control in Humans and Humanoid Robotics. *Adaptive Motion of Animals and Machines* (1): 261–280.

Schaal, S. and Atkeson, C. G. (2010). Learning control in robotics. *IEEE Robotics and Automation Magazine* **17**(2): 20–29.

Schiavi, R., Grioli, G., Sen, S. and Bicchi, A., VSA-II: a Novel Prototype of Variable Stiffness Actuator for Safe and Performing Robots Interacting with Humans. In *2008 IEEE International Conference on Robotics and Automation*. IEEE, 2008, 2171–2176.

Seok, S., Wang, A., Chuah, M. Y. M., Hyun, D. J., Lee, J., Otten, D. M., Lang, J. H. and Kim, S. (2015). Design Principles for Energy-Efficient Legged Locomotion and Implementation on the MIT Cheetah Robot. *IEEE/ASME Transactions on Mechatronics* **20**(3): 1117–1129.

Siegel, A. and Sapru, H. (2005). *Essential Neuroscience*.

Sparrow, W. and Newell, K. M. (1998). Metabolic energy expenditure and the regulation ofmovement economy. *Psychonomic Bulletin & Review* **5**(2): 173–196.

Spong, M. W. (1987). Modeling and Control of Elastic Joint Robots. *Journal of Dynamic Systems, Measurement, and Control* **109**(4): 310–318.

Spröwitz, A. T., Tuleu, A., Ajallooeian, M., Vespignani, M., Möckel, R., Eckert, P., D'Haene, M., Degrave, J., Nordmann, A., Schrauwen, B., Steil, J. and Ijspeert, A. J. (2018). Oncilla Robot: A Versatile Open-Source Quadruped Research Robot With Compliant Pantograph Legs. *Frontiers in Robotics and AI* **5**(0): 1–18. 1803.06259.

Stulp, F. and Sigaud, O. (2013). Robot Skill Learning: From Reinforcement Learning to Evolution Strategies. *Paladyn, Journal of Behavioral Robotics* **4**(1).

Stulp, F., Theodorou, E. A. and Schaal, S. (2012). Reinforcement Learning With Sequences of Motion Primitives for Robust Manipulation. *IEEE Transactions on Robotics* **28**(6): 1360–1370.

Tagliamonte, N. L., Sergi, F., Accoto, D., Carpino, G. and Guglielmelli, E. (2012). Double actuation architectures for rendering variable impedance in compliant robots: A review. *Mechatronics* **22**(8): 1187–1203.

Tassa, Y., Mansard, N. and Todorov, E., Control-limited differential dynamic programming. In *ICRA*. 2014, 1168–1175.

Theodorou, E., Buchli, J. and Schaal, S. (2010a). A Generalized Path Integral Control Approach to Reinforcement Learning. *J. Mach. Learn. Res.* **11**: 3137–3181.

Theodorou, E., Buchli, J. and Schaal, S. (2010b). Reinforcement learning of motor skills in high dimensions: A path integral approach. *2010 IEEE International Conference on Robotics and Automation* **1**(3): 2397–2403.

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature neuroscience* **7**(9): 907–15.

Todorov, E. and Jordan, M. (2002). Optimal feedback control as a theory of motor coordination. *Naturenature Neuroscience* (5): 1226–1235.

Tomei, P. (1991). A Simple PD Controller for Robots with Elastic Joints. *IEEE TRANSACTIONS ON AUTOMATIC CONTROL* **36**(10): 1208–1213.

Toussaint, M., Gienger, M. and Goerick, C., Optimization of sequential attractor-based movement for compact behaviour generation. In *2007 7th IEEE-RAS International Conference on Humanoid Robots*, 2. IEEE, 2007, 122–129.

Tucker, V. A. (1975). The Energetic Cost of Moving About: Walking and running are extremely inefficient forms of locomotion. Much greater efficiency is achieved by birds, fish—and bicyclists. *American Scientist* **63**(4): 413–419.

Uno, Y., Kawato, M. and Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement. *Biological Cybernetics* **61**(2): 89–101.

Van Ham, R., Sugar, T. G., Vanderborght, B., Hollander, K. W., Lefeber, D., Ham, R., Sugar, T. G., Vanderborght, B., Hollander, K. W. and Lefeber, D. (2009). Compliant actuator designs. *IEEE Robotics & Automation Magazine* **16**(3): 81–94.

Van Ham, R., Vanderborght, B., Van Damme, M., Verrelst, B. and Lefeber, D. (2007). MACCEPA, the mechanically adjustable compliance and controllable equilibrium position actuator: Design and implementation in a biped robot. *Rob. Auton. Syst.* **55**(10): 761–768.

Vanderborght, B., Albu-Schaeffer, A., Bicchi, A., Burdet, E., Caldwell, D., Carloni, R., Catalano, M., Ganesh, G., Garabini, M., Grebenstein, M., Grioli, G., Haddadin, S., Jafari, A., Laffranchi, M., Lefeber, D., Petit, F., Stramigioli, S., Tsagarakis, N., Van Damme, M., Van Ham, R., Visser, L. C. and Wolf, S., Variable impedance actuators: Moving the robots of tomorrow. In *IEEE International Conference on Intelligent Robots and Systems*. 2012.

Vanderborght, B., Albu-Schäffer, A., Bicchi, A., Burdet, E., Caldwell, D., Carloni, R., Catalano, M., Eiberger, O., Friedl, W., Ganesh, G., Garabini, M., Grebenstein, M., Grioli, G., Haddadin, S., Hoppner, H., Jafari, A., Laffranchi, M., Lefeber, D., Petit, F., Stramigioli, S., Tsagarakis, N., Van Damme, M., Van Ham, R., Visser, L. and Wolf, S. (2013). Variable impedance actuators: A review. *Robotics and Autonomous Systems* **61**(12): 1601–1614.

Vanderborght, B., Tsagarakis, N. G., Van Ham, R., Thorson, I. and Caldwell, D. G. (2011). MACCEPA 2.0: Compliant actuator used for energy efficient hopping robot Chobino1D. *Autonomous Robots* **31**(1): 55–65.

VanderBorght, B., Verrelst, B., Van Ham, R., Van Damme, M., Beyl, P. and Lefeber, D. (2008). Development of a compliance controller to reduce energy consumption for bipedal robots. *Autonom. Robots* **24**(4): 419–434.

Vanderborght, B., Verrelst, B., Van Ham, R., Van Damme, M., Lefeber, D., Duran, B. M. Y. and Beyl, P. (2006). Exploiting Natural Dynamics to Reduce Energy Consumption by Controlling the Compliance of Soft Actuators. *The International Journal of Robotics Research* **25**(4): 343–358.

Velasco, A., Garabini, M., Catalano, M. G. and Bicchi, A., Soft actuation in cyclic motions: Stiffness profile optimization for energy efficiency. In *IEEE-RAS International Conference on Humanoid Robots*. 2015, 107–113.

Verstraten, T., Furnemont, R., Mathijssen, G., Vanderborght, B. and Lefeber, D. (2016). Energy Consumption of Geared DC Motors in Dynamic Applications: Comparing Modeling Approaches. *IEEE Robotics and Automation Letters* **1**(1): 524–530.

Visser, L. C., Carloni, R. and Stramigioli, S., Energy Efficient Control of Robots with Variable Stiffness Actuators. In *IFAC Proceedings Volumes*, volume 43. 2010, 1199–1204.

Visser, L. C., Stramigioli, S. and Bicchi, A., Embodying desired behavior in variable stiffness actuators. In *IFAC Proceedings Volumes*, volume 18. 2011, 9733–9738.

Wada, T., Ishikawa, M., Kitayoshi, R., Maruta, I. and Sugie, T., Practical modeling and system identification of R/C servo motors. In *IEEE International Conference on Control Applications*. 2009, 1378–1383.

Wilson, J. M. and Flanagan, E. P. (2008). The Role of Elastic Energy in Activities with High Force and Power Requirements: A Brief Review. *Journal of Strength and Conditioning Research* **22**(5): 1705–1715.

Wolf, S., Bahls, T., Chalon, M., Friedl, W., Grebenstein, M., Höppner, H., Kühne, M., Lakatos, D., Mansfeld, N., Özparpucu, M. C., Petit, F., Reinecke, J., Weitschat, R. and Albu-Schäffer, A. (2015). *Soft Robotics with Variable Stiffness Actuators: Tough Robots for Soft Human Robot Interaction*. Berlin, Heidelberg, Springer Berlin Heidelberg, 231–254.

Wolf, S., Eiberger, O. and Hirzinger, G., The DLR FSJ: Energy based design of a variable stiffness joint. In *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, 5082–5089.

Wolf, S. and Hirzinger, G., A new variable stiffness design: Matching requirements of the next robot generation. In *2008 IEEE International Conference on Robotics and Automation*. IEEE, 2008, 1741–1746.

Wu, F. and Howard, M., A Hybrid Dynamic-Regenerative Damping Scheme for Energy Regeneration in Variable Impedance Actuators. In *IEEE Int. Conf. Robotics & Automation*. IEEE, 2018, 4277–4282.

Wu, F. and Howard, M. (2020a). Energy Regenerative Damping in Variable Impedance Actuators for Long-Term Robotic Deployment. *IEEE Transactions on Robotics* : 1–13.

Wu, F. and Howard, M. (2020b). Exploiting Variable Impedance for Energy Efficient Sequential Movements. *arXiv preprint* (arXiv:2002.12075).

Yamaguchi, J., Nishino, D. and Takanishi, A., Realization of dynamic biped walking varying joint stiffness using antagonistic driven joints. In *IEEE International Conference on Robotics and Automation*. 1998, 2022–2029.

Yokoi, A. and Diedrichsen, J. (2019). Neural Organization of Hierarchical Motor Sequence Representations in the Human Neocortex. *Neuron* **103**(6): 1178 – 1190.e7.

Yu, H., Huang, S., Thakor, N. V., Chen, G., Toh, S. L., Sta Cruz, M., Ghorbel, Y. and Zhu, C., A novel compact compliant actuator design for rehabilitation robots. In *IEEE International Conference on Rehabilitation Robotics*. 2013.

Yuan, Y.-x. (2015). Recent advances in trust region algorithms. *Mathematical Programming* **151**(1): 249–281.

Zhakatayev, A., Rubagotti, M. and Atakan Varol, H. (2017). Time-Optimal Control of Variable-Stiffness-Actuated Systems. *IEEE/ASME Transactions on Mechatronics* **22**(3): 1247–1258.

Zhang, R., Wang, X. and John, S. (2018). A comprehensive review of the techniques on regenerative shock absorber systems. *Energies* **11**(5).

Zhao, Y., Sena, A., Wu, F. and Howard, M. J., A Framework for Teaching Impedance Behaviours by Combining Human and Robot 'Best Practice'. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2018, 3010–3015.